PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07H 21/04, C12N 1/12, 1/14, 1/20, 15/63, C12P 21/02, C12Q 1/00, 1/02, 1/04

A1

(11) International Publication Number: WO 99/52926

(43) International Publication Date: 21 October 1999 (21.10.99)

(21) International Application Number: PCT/US99/08164

(22) International Filing Date: 14 April 1999 (14.04.99)

(30) Priority Data:

60/098,563	14 April 1998 (14.04.98)	US
60/082,952	24 April 1998 (24.04.98)	US
60/100,430	10 July 1998 (10.07.98)	US
60/105,441	23 October 1998 (23.10.98)	US
60/105,447	23 October 1998 (23.10.98)	US
60/117,758	29 January 1999 (29.01.99)	US
60/117,955	29 January 1999 (29.01.99)	US

(71) Applicant *(for all designated States except US)*: VERSICOR, INC. [US/US]; 34790 Ardentech Court, Fremont, CA 94555 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: TRIAS, Joaquim [ES/US]; Apartment 305, 407 Turner Terrace, San Mateo, CA 94401 (US). YOUNG, Dennis [US/US]; 743 Cornelia Court, Mountain View, CA 94040 (US). ROSENOW, Carsten [US/US]; 105 Livorno Way, Redwood City, CA 94065 (US).

(74) Agents: BRENNAN, Sean, M. et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304–1018 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: REGULATED TARGET EXPRESSION FOR SCREENING

(57) Abstract

Methods and compositions for screening compounds for potential therapeutic activity and for identifying drug targets are provided. The methods rely on controlled expression (either underexpression or overexpression) of an essential cellular gene, which can be achieved, in one embodiment, by fusion of a heterologous regulatory element to the gene. The method is capable of identifying a drug target in the absence of any knowledge of target function.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

REGULATED TARGET EXPRESSION FOR SCREENING

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority under 35 U.S.C. § 119(e) to the following U.S. Provisional Patent applications: 60/098,563, filed April 14, 1998; 60/082,952, filed April 24, 1998; 60/100,430, filed July 10, 1998; 60/105,441, filed October 23, 1998; 60/105,447, filed October 23, 1998; 60/117,758, filed January 29, 1999 and 60/117,955, filed January 29, 1999. The disclosures of all of these applications are hereby incorporated by reference herein in their entireties.

TECHNICAL FIELD

This invention is in the field of drug screening and drug discovery. More particularly, techniques of microbial genetics are utilized to provide methods and compositions for identifying targets and screening candidate therapeutics.

BACKGROUND

Many methods exist for the discovery of novel therapeutic agents, such as antibiotics. Cell-free, target-based assays often identify potent target inhibitors, but inhibitors identified in this fashion often exhibit no activity or only poor activity against whole cells. *See*, for example, Isaacson (1994) *Exp. Opin. Investig. Drugs* **3**:83-91; J. Sutcliffe and N. Georgopapadakou (*eds.*) Emerging Targets in Antibacterial and Antifungal Chemistry, Chapman and Hall, London, 1992. Whole-cell methods have traditionally involved screening compounds against wild-type strains of pathogens and selecting as candidates those compounds which have a negative effect on the viability of the pathogen. Under

these conditions, compounds selected as candidates generally interact with a target that is expressed at wild-type levels. The potential of this type of assay is limited, since it provides no information on mechanism of action, which is critical for selection of a candidate. Identification of drug targets and determination of target function are costly and time-consuming processes. One approach to overcoming some of these problems has been to isolate mutants that are hypersusceptible to a particular agent and use them to screen for new agents having similar properties and/or mechanisms of action. In addition, compounds having activity against a hypersusceptible strain can often, with minimal modification, be converted to agents with strong activity against the wild-type strain. Such hypersusceptible mutants are generally obtained following standard chemical mutagenesis with agents such as N-methyl-N'-nitro-N-nitrosoguanidine. *See*, for example, Kitano *et al.* (1977) *The Japanese Journal of Antibiotics*, vol. XXX Suppl., pp. S239-S245; Numata *et al.* (1986) *The Journal of Antibiotics*, vol. XXXIX, pp. 994-1000; and Kamogashira *et al.* (1988) *The Journal of Antibiotics*, vol. XLI, pp. 803-806.

Systems for regulated expression of cloned genes have been described. These include the following promoters: *trp, lpp, lac, tac, trc,* λP_L, λP_R, *tet*A, *rec*A, *pho*A, *mal*X, *mal*M (*S. pneumoniae*), *xyl (S. carnosus)* and T7. *See*, for example, Tacon *et al.* (1980) *Mol. Gen. Genet.* **177**:427-438; Ghrayeb *et al.* (1984) *EMBO J.***3**:2437-2442; Germino *et al.* (1983) *Cell* **32**:131-140; Russell *et al.* (1982) *Gene* **20**:231-243; Hallewell *et al.* (1985) *Nucleic Acids Res.* **13**:2017-2034; Yoakum *et al.* (1982) *Proc. Natl. Acad. Sci. USA* **82**: 1766-1770; Queen (1983) *J. Mol. Appl. Genet.* **2**:1-10; De la Torre *et al.* (1984) *J. Biol. Chem.* **259**:11571-11575; Shirakawa *et al.* (1984) *Gene* **28**:127-132; Miyake *et al.* (1985) *J. Biochem.* **97**:1429-1436; Studier *et al.* (1986) *J. Mol. Biol.* **189**:113-130; Johnston *et al.* (1985) *Gene* **34**:137-145; Nieto *et al.* (1997) *J. Biol. Chem.* **272**:30860-30865; and Sizemore *et al.* (1993) *FEMS Microbiol. Lett.* **107**:303-306. For general reviews, see Bauerle (ed.) "Inducible Gene Expression" Birkhauser, Boston, 1985; A. Smith (ed.) "Gene expression in recombinant microorganisms," M. Dekker, New York, 1994; Makrides (1996) *Microbiol. Rev.*

60:512-538; and de Vos *et al.* (1997) *Curr. Opin. Biotechnol.* **8**:547-553. Most of the above-mentioned systems are capable of overexpression of one or more cloned genes. These systems often exhibit moderate-to-high basal expression levels, above which overexpression can be induced by manipulation of environmental conditions and/or provision of inducing molecules. Fusions between the P_{BAD} promoter of the arabinose operon and a heterologous gene have been used for the overexpression of heterologous genes. *See* U.S. Patent No. 5,028,530. However, in contrast to other systems, the P_{BAD} promoter can be extremely tightly regulated to provide very low basal levels of expression. *See*, for example, Guzman *et al.* (1995) *J. Bacteriology* **177**:4121-4130. Construction of arabinose-dependent strains, generated by placing an essential gene under the control of *ara* regulatory elements, has been described. *See*, for example, Brown *et al.* (1995) *J. Bacteriol.* **177**:4194-4197; Dalbey *et al.* (1985) *J. Biol. Chem.* **260**:15925-15931; and Guzman *et al.*, *supra*.

Systems for drug screening have been described wherein overexpression of a target gene product results in acquisition of resistance to an inhibitor, identifying the gene product as a potential target of the inhibitor. *See*, for example, del Castillo *et al.* (1991) *Proc. Natl. Acad. Sci. USA* **88**:8860-8864. Screening for inhibitors of a particular enzyme has been accomplished by comparing the effect of a test compound on a strain that is defective for the enzyme with the effect of the test compound on a strain harboring a different mutation. *See*, for example, EP 644268. Other screening systems have been developed which depend on generating strains which express mutant proteins (*e.g.*, temperature-sensitive proteins) and assessing their sensitivity to test compounds. *See*, for example, PCT Publication WO 96/23075.

Accordingly, highly-inducible regulatory systems with low basal expression levels would be extremely useful for the identification of essential genes and inhibitors of essential genes in microorganisms.

DISCLOSURE OF THE INVENTION

One object of the invention is to provide methods and compositions for identifying compounds which inhibit the growth or viability of an organism, regardless of whether the mechanism of action of the inhibitor and/or the function of the inhibitor's target is known. Another object of the invention is to provide compositions and methods for screening compounds using cells that are hypersusceptible to an inhibitor. An additional object of the invention is to provide methods for generating cells that are hypersusceptible to a known inhibitor, utilizing techniques of molecular genetics and recombinant DNA, in particular, techniques that permit regulated expression of a target gene, including underexpression, expression at normal levels, and overexpression. A further object of the invention is to provide compositions and methods for identifying essential genes and gene products of microorganisms, as well as genes and gene products that are involved in virulence and drug resistance. Yet another object of the invention is to provide methods and compositions for determining the mechanism of action of an inhibitor. A further object of the invention is to provide methods and compositions for controlled gene expression. An additional object is to provide methods and compositions that will allow expression of a particular target gene to be regulated at levels that are both lower and higher than those normally present in the cell.

Accordingly, in one aspect the invention provides cells in which the expression of a gene product involved in an essential function can be regulated. In particular, the invention provides cells in which the expression of a gene can be down-regulated to express the gene product below wild-type levels, as well as cells in which gene expression can be up-regulated to levels that are higher than wild-type. In some cases, expression of a gene product at lower-than-normal levels will, in and of itself, result in an impairment or absence of growth which defines the gene product as being essential. In other cases, environmental conditions (such as, for example, temperature, pH, nutrient sources, ionic strength, presence of other organisms, infection and/or presence of a compound) under

which expression of a particular level of a given gene product is essential can be determined.

In one aspect, the methods and compositions of the invention will allow expression of any gene in a cell to be independently regulated by external stimuli, such as nutrient concentration. Genes whose expression can be down-regulated to a point at which levels of that particular gene product become limiting for growth or other important cell function (*e.g.*, pathogenesis or resistance to antibiotics) can then be identified. Once such a gene has been identified, cells expressing that gene at any level between that which is limiting and any higher expression level can be challenged with a test compound. Compounds which exhibit higher potency against cells expressing lower levels of gene product are candidate inhibitors. It can be seen that, since this method depends simply on regulating levels of a particular gene product, it is not necessary beforehand to know the function of the gene product that is being regulated, nor is it necessary to know the mechanism of action of the inhibitor. Accordingly, knowledge of target function is not necessary for the identification of an inhibitor in the practice of the invention.

In another aspect, the invention provides methods and compositions for the identification of compounds that affect essential cellular processes, by exposing to a test compound cells in which expression of a gene product that is involved in an essential cellular process is regulated to a lower-than-normal level.

In yet another aspect, the invention provides methods and compositions for determining the target and mechanism of action of an inhibitor by exposing, to a test compound, cells in which expression of a gene product that is involved in an essential cellular process is regulated to a lower-than-normal level.

In yet another aspect, the invention provides methods and compositions for determining the target and mechanism of action of an inhibitor by exposing, to a test compound, a library of cells in which expression of a variety of gene products that are involved essential cellular processes are regulated to lower-than-normal levels.

In a further aspect, the practice of the invention will allow identification of genes encoding drug targets, genes encoding essential cellular functions, genes encoding virulence factors, genes encoding antibiotic resistance factors, polypeptides or fragments thereof that serve as drug targets or virulence factors; polypeptides or fragments thereof that participate in essential cellular functions or antibiotic resistance; RNAs that serve as drug targets or virulence factors and RNAs that participate in essential cellular functions or antibiotic resistance.

Cells in which expression of a gene product that is involved in an essential cellular process is regulated to a lower-than-normal level can be obtained through techniques of microbial genetics and molecular biology. For example, fusion of a heterologous regulatory element to an essential gene places that gene under the control of the heterologous regulatory element. The heterologous regulatory element may intrinsically provide lower expression levels than the essential gene's normal regulatory system, or the heterologous regulatory element may be capable of being down-regulated. In either case, expression of the essential gene at lower-than-normal levels is possible.

The invention also provides methods and compositions for regulated gene expression, whereby expression is controlled over a range of levels ranging from underexpression through normal expression levels through overexpression. Exemplary compositions include regulatable promoters, enhancers, operators and other transcriptional and/or translational control elements. Exemplary methods include methods for placing regulatable promoters, enhancers, operators and other transcriptional and/or translational control elements into operative linkage with a gene or coding sequence, and expression of such constructs in a cell, wherein expression is regulated by an inducer and/or repressor.

Methods and compositions for regulated expression of a gene in a microorganism are also provided; wherein the methods utilize a construct comprising a gene, or a fragment thereof, in operative linkage with a regulatory element such as the *E. coli* P_{BAD} promoter or the P_{AGA} promoter of *S. pneumoniae*. The methods comprise introducing the construct into a host cell, culturing the host cell in a growth medium and adjusting the concentration of one or more

modulator substances in the growth medium. Modulator substances can be inducers and/or negative modulators (*i.e.*, repressors) of the *raf* regulatory element(s) present in the construct.

In another embodiment, compositions and methods for making a construct, comprising a gene or a fragment thereof in operative linkage with a component of the *raf* regulatory region of *S. pneumoniae*, are provided. Such constructs can be chromosomal or extrachromosomal.

The invention will therefore be useful for drug screening, target identification, determining mechanisms of action of antibiotics, determining mechanisms of virulence and antibiotic resistance, and for other purposes as will be apparent to those of skill in the art.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the nucleotide sequence of the *S. pneumoniae Px* and *Pm* regions. (SEQ ID NO: 1). The mutation in the repressor binding site of the *Px* promoter, converting a GGA sequence to GCG (mutant construction described in Example 6) is indicated on the bottom line of the figure.

Figure 2 is a schematic diagram of the *S. pneumoniae raf* gene cluster, including the two *raf* operons, with open reading frames (ORFs) represented by arrows. Locations of promoters is also indicated. Shown for comparison is a schematic diagram of ORFs in the *msm* region of *S. mutans*.

Figure 3 shows the nucleotide sequence of the *raf* region of *S. pneumoniae* strain VSPN3026 (SEQ ID NO. 2). The general location of the P_{AGA} promoter is indicated by underlining.

Figure 4 shows idealized results of an experiment in which the minimum inhibitory concentration of a compound is determined as a function of inducer concentration, in a cell in which target expression level is regulated by inducer concentration and the target is a single component which is inhibited by the compound.

Figure 5 shows idealized results of an experiment in which the minimum inhibitory concentration of a compound is determined as a function of inducer

concentration, in a cell in which either 1) the target comprises multiple components and the compound interacts with a site defined by two or more of the components or 2) the compound interacts with multiple targets, and the level of one of the components (or targets) is regulated by inducer concentration.

Figure 6 shows a scheme for replacement of wild type *mur*A regulatory elements with an *ara* regulatory cassette.

Figure 7 shows growth of a P_{BAD}-*mur*A fusion strain (*E. coli* VECO2055) as a function of arabinose concentration.

Figure 8 shows the minimum inhibitory concentration of fosfomycin, as a function of arabinose concentration, for the P_{BAD}-*mur*A fusion strain *E. coli* VECO2055, compared to wild-type.

Figure 9 shows minimum inhibitory concentrations of fosfomycin, ciprofloxacin, and tetracycline for the *E. coli* P_{BAD}-*mur*A fusion strain VECO2055, expressed as a function of arabinose concentration.

Figure 10 shows optical density measurements, at various times after inoculation, of cultures of VSPN3041 grown at different raffinose concentrations.

Figure 11 shows optical density measurements of the growth of VSPN3041 on either sucrose or raffinose.

Figure 12 shows the growth of VSPN3041 and the parent isogenic strain VSPN3026, at different raffinose concentrations. Growth was measured by optical density after 10 hours of culture.

Figure 13 shows the susceptibility of VECO2065, an *E. coli* strain having a chromosomal P_{BAD}-*def* fusion, to VRC483, an inhibitor of the *def* gene product. Susceptibility is presented as minimum inhibitory concentration (MIC) of VRC483 in μg/ml, as a function of inducer (arabinose) concentration. Also shown is the susceptibility of VECO2065 to fosfomycin and ciprofloxacin.

Figure 14 shows the susceptibility of VECO2079, an *E. coli* strain having a chromosomal P_{BAD}-*folA* fusion, to trimethoprim, an inhibitor of the *folA* gene product. Susceptibility is presented as minimum inhibitory concentration (MIC) of trimethoprim in μg/ml, as a function of inducer (arabinose) concentration. Also shown is the susceptibility of VECO2079 to fosfomycin and ciprofloxacin, and

the susceptibility of the parent strain, VECO2054 (indicated by wt), to trimethoprim.

Figure 15 shows the susceptibility of VECO2083, an *E. coli* strain having a chromosomal P_{BAD}-*gyrB* fusion, to novobiocin, an inhibitor of the *gyrB* gene product. Susceptibility is presented as minimum inhibitory concentration (MIC) of novobiocin in μg/ml, as a function of inducer (arabinose) concentration. Also shown is the susceptibility of VECO2083 to fosfomycin and ciprofloxacin.

Figure 16 shows the susceptibility of VECO2068, an *E. coli* strain having a chromosomal P_{BAD}-*def* fusion and a *tolC* deletion, to VRC483, an inhibitor of the *def* gene product (indicated by the curve labeled VRC483). Susceptibility is presented as minimum inhibitory concentration (MIC) of VRC483 in μg/ml, as a function of inducer (arabinose) concentration. Also shown is the susceptibility of VECO2068 to fosfomycin and ciprofloxacin, and the susceptibility to VRC483 of the parent strain, VECO2066 (indicated by the curve labeled VRC483, *tolC*).

Figure 17 shows the susceptibility of VSPN3044 to VRC483, an inhibitor of the *def* gene product. VSPN3044 contains a P_{AGA}-*def* transcriptional fusion, so that expression of the *def* gene product is regulated by raffinose. Susceptibility is presented as minimum inhibitory concentration (MIC) of VRC483 in μg/ml, as a function of inducer (raffinose) concentration. Also shown in the susceptibility of VSPN3044 to erythromycin and vancomycin, and the susceptibility of the parent strain VSPN3026 (indicated by VRC483wt) to VRC483.

Figure 18 shows the susceptibility of VECO2524 (P_{BAD}-*lpxC*, Δ*tolC*) toL159692, an antibacterial compound that targets the *lpxC* gene product. Minimum inhibitory concentration is shown as a function of arabinose concentration. Also shown are minimum inhibitory concentrations of linezolid and erythromycin as a function of arabinose concentration.

MODES FOR CARRYING OUT THE INVENTION

The practice of the present invention will employ, unless otherwise indicated, conventional techniques in organic chemistry, biochemistry, molecular biology, microbiology, genetics, recombinant DNA, and related fields as are

within the skill of the art. These techniques are fully explained in the literature. See, for example, Maniatis, Fritsch & Sambrook, MOLECULAR CLONING: A LABORATORY MANUAL, Cold Spring Harbor Laboratory Press (1982); Sambrook, Fritsch & Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, Second Edition, Cold Spring Harbor Laboratory Press (1989); Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons (1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996); Silhavy *et al.*, EXPERIMENTS WITH GENE FUSIONS, Cold Spring Harbor Laboratory Press (1984); Gerhardt *et al.*, METHODS FOR GENERAL AND MOLECULAR MICROBIOLOGY, American Society for Microbiology, Washington, D.C., 1994; Lorian, ANTIBIOTICS IN LABORATORY MEDICINE, 4th ed., Williams & Wilkins, Baltimore, 1996; and Murray *et al.* MANUAL OF CLINICAL MICROBIOLOGY, 6th ed., American Society for Microbiology, Washington, D.C., 1995.

All patents, patent applications and publications cited herein are incorporated by reference in their entirety.

The present invention provides methods and compositions useful for identification of compounds that affect an essential cellular process, compounds that interfere with mechanisms of resistance, and compounds that interfere with virulence factors; for identification of the target or targets of a compound that affects an essential cellular process, a mechanism of resistance or a virulence factor; for identification of a gene or genes encoding a target or targets of a compound that affects an essential cellular process, a mechanism of resistance or a virulence factor; and for identification of genes, RNAs and polypeptides involved in essential cellular processes, mechanisms of resistance or virulence. Identification is facilitated by controlled expression of a gene that is involved in an essential cellular process. Knowledge of the function of a gene or its product is not required, either to identify it as being involved in an essential cellular process, or to identify a compound which affects the gene product.

Cell types

In one embodiment, the present invention will be used in the identification of compounds which have activity against microorganisms. Accordingly, compositions embodied by the invention will include microorganisms wherein the expression of an essential gene of the microorganism is regulated by fusion to a heterologous regulatory element. Similarly, target genes and polypeptides whose expression is regulated by a heterologous regulatory element will often be those that are essential for viability of a microorganism, or responsible for its virulence or drug resistance.

Microorganisms can be either prokaryotic or eukaryotic; and prokaryotes can be either Gram-positive or Gram-negative. Exemplary prokaryotes include, but are not limited to: *Staphylococcus (e.g., S. aureus, S. epidermidis), Streptococcus (e.g., S. pneumoniae, S. pyogenes, S. agalactiae), Enterococcus (E. faecalis, E. faecium), Neisseria, Branhamella, Listeria, Bacillus (e.g., B. subtilis), Corynbacterium, Erysipelothrix, Gardnerella, Nocardia, Mycobacterium,* enterobacteriaceae, *Escherichia (e.g., E. coli), Salmonella, Shigella, Yersinia, Enterobacter (e.g., E. cloacae), Klebsiella (e.g., K. pneumoniae, K. oxytoca), Citrobacter, Serratia, Providencia, Proteus (e.g., P. mirabilis, P. vulgaris), Morganella (e.g., M. morganii), Edwardsiella, Erwinia, Vibrio, Aeromonas, Helicobacter (e.g., H. pylori), Campylobacter, Eikenella, Pasteurella, Pseudomonas (e.g., P. aeruginosa), Burkholderia, Stenotrophomonas, Acinetobacter, Ralstonia, Alcaligenes, Moraxella, Legionella, Francisella, Brucella, Haemophilus (e.g., H. influenzae), Bordetella, Clostridium, Bacteroides, Porphyromonas, Prevotella, Fusobacterium, Borrelia, Chlamydia, Ricketsia, Ehrlichia* and *Bartonella.*

Exemplary eukaryotic microorganisms include, but are not limited to, yeasts and fungi, for example, *Candida (e.g., C. albicans), Cryptococcus, Pneumocystis, Histoplasma, Blastomyces, Coccidioides, Aspergillus, Fusarium, Saccharomyces* and *Schizosaccharomyces.*

The practice of the invention can also be applied to eukaryotic cells, such as plant cells, mammalian cells and human cells. In one embodiment, malignant

cells which are resistant to a therapeutic can be analyzed to determine the locus of resistance and to identify compounds that will reverse resistance by interacting with the cellular component responsible for resistance. In this context, a therapeutic can comprise a compound, such as a drug, a composition comprising multiple compounds, or a physical treatment, such as radiation.

Essential cellular functions

In one embodiment, the invention provides methods and compositions for identifying genes and/or gene products involved in essential cellular functions. An essential function for a particular cell will depend on the genotype of the cell and the cell's environment. By way of example, essential cellular functions are those which are involved in replication, repair, recombination and transcription of genetic material; protein synthesis (translation), processing and transport; protein export; anabolic synthesis of cellular molecules; catabolism of cellular nutrients; synthesis of cell membranes and cell walls; lipid metabolism; protein metabolism; energy metabolism; cell division; cell shape; filamentation; regulation; DNA binding; RNA binding; efflux systems; transport systems; virulence or pathogenicity; and drug resistance. Protein metabolism can include protein modifications such as glycosylation, phosphorylation, acetylation and ubiquitination, to name but a few examples. Gene products that can be involved in essential cellular processes include, but are not limited to, topoisomerases, nucleases, recombinases, primases, helicases, DNA polymerases, RNA polymerases, histone modifying enzymes, kinases, phosphatases, acetylases, deacetylases, formylases, deformylases, chaperonins, ion transporters, cytoskeletal elements, colicins, cytochromes, ribosomal proteins, transfer RNAs, ribosomal RNAs, hydrolases, proteases, epimerases, rotamases, synthases, racemases, dehydrogenases, transferases, ligases, reductases, oxidases, transglycosylases, transpeptidases, peptidases, GTPases, ATPases, translocases, ribonucleases, transcription factors, sigma factors, ribosomal release factors, structural RNAs and structural proteins.

More generally, an essential cellular process is any process which, when it occurs at a lower rate or to a lesser extent than normal, negatively influences the viability of the cell. Methods for determination of cell viability are well-known to those of skill in the art and include, but are not limited to, vital staining, cell counting, either microscopically or by colony counting following serial dilution and plating of cell cultures, measurement of light scattering by cell cultures, fluorescence-activated cell sorting, incorporation of polynucleotide and/or polypeptide precursors, reporter gene expression, and measurement of cell weight and/or volume.

The types of molecules that can participate in essential cellular processes can include nucleic acids, polypeptides and other cellular macromolecules. Nucleic acids will include, for example, DNA; regulatory RNA molecules, such as ribozymes and antisense RNA; transfer RNA and ribosomal RNA. Polypeptides can include, for example, structural proteins, enzymes, receptors, intracellular signaling molecules, and cellular adhesion molecules.

Regulatory elements

In one aspect of the invention, the expression of a gene involved in an essential cellular function is regulated by fusion of the gene, or a fragment thereof, to a heterologous regulatory element. A heterologous regulatory element is one that is not normally associated with, and does not normally regulate, the gene which it regulates in the practice of the invention. Regulatory elements can comprise transcriptional, post-transcriptional, translational, and post-translational elements; as well as regulatory elements related to replication. By way of example, transcriptional regulatory elements can include promoters, enhancers, operators, and elements that modulate the rate of transcription initiation, elongation and/or termination; post-transcriptional regulatory elements can include those influencing messenger stability, processing and transport; translational regulatory elements can include those which modulate the frequency of translation initiation and the rate of translational elongation; post-translational regulatory elements can include those which influence protein processing,

stability and transport; and replication-associated regulatory elements can include those related to gene dosage.

In preferred embodiments, the heterologous regulatory element comprises a regulatable promoter. In a particularly preferred embodiment, the regulatable promoter is the *ara*BAD promoter, also known as P_{BAD}. Regulation by P_{BAD} has been the subject of extensive study and its regulatory properties are well-understood. *See*, for example, Schleif (1992) *Ann. Rev. Biochem.* **61**:199-223; Guzman *et al.* (1995) *J. Bacteriology* **177**:4121-4130; and Gallegos *et al.* (1997) *Microbiology and Molecular Biology Reviews* **61**:393-410.

The P_{BAD} promoter is regulated by the AraC protein, which has both positive and negative regulatory activities. In the absence of L-arabinose or other inducers (such as, for example, L-ribose), AraC represses transcription from P_{BAD} by binding to sites upstream of the P_{BAD} transcription initiation site. Inducers such as L-arabinose interact with the Ara*C* protein to form an activator of P_{BAD} transcription that binds to different upstream sites to stimulate transcription. With respect to the present invention, desirable features of the AraC/P_{BAD} regulatory system are the very low basal levels of transcription obtained in the absence of arabinose and the direct relationship between transcription from P_{BAD} and the concentration of arabinose in the medium. *See*, for example, Guzman *et al.*, *supra*.

The activity of P_{BAD} is directly proportional to the concentration of arabinose in the environment and, importantly, at low arabinose concentration, very low basal levels of expression are obtained. The P_{BAD} promoter is also subject to regulation by catabolite repression, mediated by cyclic AMP and by the cyclic AMP receptor protein, also known as the catabolite repressor protein (CRP). Thus, further modulation of P_{BAD} expression can be obtained by regulating the concentration of glucose (or other carbon source such as, for example, glucose-6-phosphate) in the environment, which modulates CRP activity within the cell. In particular, minimal expression of P_{BAD} (maximal repression) is obtained in the presence of glucose and the absence of arabinose. Withdrawal of glucose from the medium and addition of arabinose (or another inducer) results in

rapid induction of transcription from P_{BAD} wherein the expression level is proportional to the arabinose concentration. Expression levels varying over a 1,000-fold range can be obtained, depending on the inducer concentration. *See*, for example, Guzman *et al.*, *supra*.

The P_{BAD} promoter or any other promoter of the AraC/XylS family, from any prokaryotic or eukaryotic organism, can be used in the practice of the invention. *See*, for example, Gallegos *et al.*, *supra*; de Vos *et al.* (1997) *Curr. Opin. Biotechnol.* **8**:547-553; and Kleerebezem *et al.* (1997) *Mol. Microbiol.* **24**:895-904. Particularly preferred are the P_{BAD} promoters of *E. coli* and *S. typhimurium*.

Another regulatory system that is useful in the practice of the invention is the *malM/malX* system of *S. pneumoniae*, regulated by MalR. MalR is a repressor that controls the expression of the maltosaccharide regulon in *S. pneumoniae* and belongs to the LacI-GalR family of repressors. Two operons are regulated in opposite direction, *malXCD* (*Px* promoter) and *malMP (Pm* promoter), see Figure 1 (SEQ ID NO: 1). Stassi *et al.* (1982) *Gene* **20**:359-366; and Nieto *et al.* (1997) *J. Biol. Chem.* **272**:30860-30865. Affinity of MalR for *Pm* is higher than for *Px* and, in both cases, a high basal level of expression has been reported. Nieto *et al.* (1997) *J. Biol. Chem.* **272**:30860-30865. Example 6, *infra*, describes fusion of *mal Px* to a catalase gene and modification of the *mal Px* promoter to obtain tight regulation by maltose in minimal medium.

Yet another example of a regulatory system that is useful in the practice of the invention is the *raf* regulatory system of *Streptococcus pneumoniae*. Example 14 shows that the *rafR* gene product acts as a positive regulator of promoters such as P_{AGA}, the promoter for the *S. pneumoniae* α-galactosidase gene. Thus, fusion of a target gene to P_{AGA}, in a cell expressing *rafR* function, will allow raffinose-regulated expression of the target gene. *See* Example 7. Additional regulatory elements in the *raf* regulatory system include the promoter of the *rafR* gene, P_{rafR}, and the promoter of the *rafE* gene, P_{rafE}.

The P_{AGA} promoter was discovered through a search of the *S. pneumoniae* genome sequence, disclosed at http://www.tigr.org. The sequence was searched

for sequences that might encode homologues to the AraC/XylS family of transcriptional activators. Gallegos *et al.* (1997) *Microbiol. Molec. Biol. Reviews* **61**:393-410. An open reading frame (ORF) encoding a protein homologous to the *Streptococcus mutans msmR* gene was identified and named *rafR*. The organization of additional ORFs in the vicinity of *rafR* was also investigated. As a result of these investigations, a gene cluster was identified, comprising two ORFs encoding regulatory proteins, *rafR* and *rafS*, an intergenic region, and six ORFs encoding structural proteins. *See* Figure 2. The gene cluster contains two operons transcribed divergently: a regulatory operon encoding *rafR* and *rafS*, and a metabolic operon which encodes *aga*, *rafE*, *rafF*, *rafG*, *gtfA* and possibly *rafH*.

The nucleotide sequence of the region of the *raf* gene cluster encompassing *rafS*, *rafR*, and *aga*, in *S. pneumoniae* strain VSPN3026 was determined. See Figure 3 (SEQ ID NO: 2). The *rafS* gene was determined to lie between the complements of nucleotide coordinates 1001-291 of this sequence, with the region encoding RafS protein lying between the complements of nucleotides 938-294. The *rafR* gene was determined to lie between the complements of nucleotides 1798-935, with the RafR coding region complementary to nucleotides 1795-938. The *aga* gene extended from nucleotides 1903-4065, with the coding region lying between 1903-4062. Several differences between the *raf* sequences of VSPN3026 (Figure 3) and those disclosed in the database at http://www.tigr.org were detected. These differences, presented in Table 1, are likely to represent polymorphisms between different strains of *S. pneumoniae*.

Table 1: Unique sequences in the *S. pneumoniae raf* region of VSPN 3026

Position in VSPN3026 sequence (SEQ ID NO: 2)	SEQUENCE	
	VSPN 3026	database
326-329	ATCC	ATACC
441	A	G
561	T	A
633	A	G
765	T	C
794	T	C
828	A	G
842	A	C
953	C	T
997	A	C
1490	A	G
1513	A	C
1665	C	G
1760	A	G
1792	G	A
2157	T	C
2739	C	T
2844	T	G
3191-3192	AT	GC
3287	C	T
3297	G	T
3399	A	G
3405	C	T
3495	A	G
3662	A	T
3693	G	A
3818	C	T

The *S. pneumoniae raf* gene cluster is organized into two domains. One domain includes the two regulatory ORFs *raf R* and *rafS*, and the other includes genes that are probably involved in uptake and catabolism, based on their

homology to *S. mutans* genes. *See* Figure 2. The term "*raf* gene cluster" refers to the *raf* transcriptional units and their related regulatory genes, in particular the region of the *S. pneumoniae* genome comprising the *rafR*, *rafS*, *aga*, *rafE*, *rafF*, *rafG*, *gtfA* and *rafH* genes, as well as the intergenic regions associated with these genes. Intergenic regions refer to DNA sequences which do not encode protein, but which lie adjacent to protein-coding regions of DNA sequence. Intergenic regions will often contain regulatory sequences such as promoters and operators, although regulatory sequences can also be located in coding regions.

Directly upstream of *rafR* is a divergently-transcribed gene, *aga*, with sequence homology to *S. mutans* α-galactosidase. Construction of a strain with a mutation in the *aga* region, followed by α-galactosidase assay of the mutant strain, shows that the *S. pneumoniae aga* gene does indeed encode a polypeptide with α-galactosidase activity. *See* Example 14. Downstream of *aga* are additional genes encoding proteins homologous to the *msm* transport system, and a gene called *gtfA*, which is a homologue of *S. mutans* sucrose phosphorylase. Although it contains several homologous ORFs, the fact that the *S. pneumoniae raf* gene cluster contains two regulatory genes suggests that its regulation may be more complex than that of the *msm* gene cluster in *S. mutans*.

The *S. pneumoniae raf* gene cluster contains at least two regulatory genes, *rafR* and *rafS*, an intergenic region, and at least five structural genes: *aga, rafE, rafF, rafG* and *gtfA*. *See* Figure 2. Sequences which regulate the expression of the regulatory and structural genes of the *S. pneumoniae raf* gene cluster are likely to be found in the intergenic region and within genes adjacent to the intergenic region. Such sequences are denoted *raf* regulatory sequences and include, for instance, promoter and operator sequences, such as the *rafR* promoter (P_{rafR}), the α-galactosidase promoter (P_{AGA}) and the *rafE* promoter P_{rafE}. Promoter sequences are those to which RNA polymerase binds to initiate transcription. Operator sequences are those to which regulatory proteins (such as, for example, activators and repressors) bind, thereby influencing the ability of RNA polymerase to bind to the promoter. In general, repressors inhibit binding of RNA polymerase, and activators facilitate binding (or relieve repressor-mediated inhibition). Transcript

analysis by RT-PCR has provided results consistent with the locations of the P_{AGA}, P_{rafE} and P_{rafR} promoters being as shown in Figure 2. Based on the presence of sequence elements homologous to well-known prokaryotic transcriptional regulatory sequences, the location of P_{AGA} was determined to be between nucleotides 1796-1902 of the sequence presented in Figure 3. The P_{rafR} promoter is also believed to lie within this region.

In one embodiment, the invention provides sequences from the *raf* regulatory region, such as *raf* promoter and operator sequences, for the regulated expression of coding sequences, which can include, for example, homologous and heterologous genes or gene fragments. With respect to the *S. pneumoniae* regulatory sequences disclosed herein, a homologous gene is one that is normally found in association with the regulatory sequences in nature. A heterologous sequence, by contrast, is a sequence from *S. pneumoniae* or any other organism, that is not normally found in association with *S. pneumoniae raf* regulatory sequences in nature. Exemplary *S. pneumoniae* regulatory sequences include, but are not limited to, the *rafR* promoter (P_{rafR}), the α-galactosidase promoter (P_{AGA}), and the promoter of the *rafE* gene, P_{rafE}.

As discussed *supra* with respect to sequence homology, and demonstrated experimentally in Example 14 *infra*, the *rafR* gene product acts as a positive regulator and the *rafS* gene product acts as a negative regulator of the *raf* operons. Growth of cells in the presence of raffinose induces expression of genes under the control of *raf* regulatory sequences, while growth of cells on sugars other than raffinose inhibits expression of genes under the control of *raf* regulatory sequences. Consequently, the methods and compositions provided by the invention allow for both overexpression and underexpression of a gene, mediated by *raf* regulatory sequences. The basal level of expression is low and the range of expression level between repressed (cells grown on maltose, for example) and induced (cells grown on raffinose) conditions is approximately a thousand-fold. *See* Example 14, *infra*.

In one embodiment, the invention provides recombinant constructs for regulation of expression of a gene of interest. The recombinant constructs are

made using standard methods of molecular biology and biotechnology to place a coding sequence in operative linkage with *raf* regulatory region sequences, either by insertion of a coding sequence in proximity to a *raf* regulatory sequence, or by insertion of a *raf* regulatory sequence in proximity to coding sequence. In preferred embodiments, the *raf* regulatory sequence will be upstream of the coding sequence when they are placed in operative linkage. Locations of restriction enzyme recognition sequences within the *raf* gene cluster, for use as insertion sites, can be easily determined by one of skill in the art from the nucleotide sequence of the *raf* gene cluster. Alternatively, various *in vitro* techniques can be used for insertion of a restriction enzyme recognition sequence at a particular site, or for insertion of heterologous sequences at a site that does not contain a restriction enzyme recognition sequence. Such methods include, but are not limited to, oligonucleotide-mediated heteroduplex formation for insertion of one or more restriction enzyme recognition sequences (*see*, for example, Zoller *et al.* (1982) *Nucleic Acids Res.* **10**:6487-6500; Brennan *et al.* (1990) *Roux's Arch. Dev. Biol.* **199**:89-96; and Kunkel *et al.* (1987) *Meth. Enzymology* **154**:367-382) and PCR-mediated methods for insertion of longer sequences. *See*, for example, Zheng *et al.* (1994) *Virus Research* **31**:163-186.

Operative linkage refers to an arrangement of one or more regulatory sequences with one or more coding sequences, such that the regulatory sequence(s) is capable of exerting its regulatory effect on the coding sequence. By way of illustration, a transcriptional regulatory sequence or a promoter is operably linked to a coding sequence if the transcriptional regulatory sequence or promoter promotes transcription of the coding sequence. Similarly, an operator is considered operatively linked to a promoter or to a coding sequence if binding of a repressor to the operator inhibits initiation at the promoter so as to prevent or diminish expression of the coding sequence. An operably linked transcriptional regulatory sequence is generally joined in *cis* with the coding sequence, but it is not necessarily directly adjacent to it.

Recombinant constructs comprising coding sequences in operative linkage with one or more *raf* regulatory region sequences can also comprise other types of

sequence including, but not limited to, replication origins, selectable markers (including, but not limited to, those encoding antibiotic resistance), transcription termination sites, sequences specifying translation initiation and termination, sequences mediating mRNA processing and/or stability and multiple cloning sites. In preferred embodiments, these additional sequences are functional in Gram-positive microorganism, such as, for example, *Streptococci*, *Staphylococci*, *Enterococci, and Lactococci*. Preferred species include, for example, *S. pneumoniae, S. pyogenes, S. agalactiae*, Lancefield group A streptococci, Lancefield group B streptococci, Lancefield group C streptococci, Lancefield group F streptococci, Lancefield group G streptococci, and viridans streptococci. Preferred non-streptococcal species in which these additional sequences are functional include, for example, enterococci such as *E. faecalis,* and *E. faecium*, and lactococci such as *L. lactis*. Methods for the construction of such recombinant constructs are well-known to those of skill in the art. *See*, for example, Sambrook *et al, supra*. It will also often be useful to include a selectable marker in the recombinant construct, to aid in the isolation and identification of cells comprising the construct. Selectable markers include those which facilitate positive selection, such as a sequence which encodes antibiotic resistance, and those which facilitate negative selection. Bochner *et al.* (1980) *J. Bacteriol.* **143**:926-933; and Gay *et al.* (1985) *J. Bacteriol.* **164**:918-921. Recombinant constructs can exist as freely-replicating extrachromosomal elements, such as plasmids or episomes, or can exist as chromosomal recombinants, such as would be achieved either by integration of a *raf* regulatory cassette into the chromosome of a microorganism adjacent to a gene of interest, or by insertion of a gene of interest into the chromosome adjacent to a *raf* regulatory sequence, for example. Methods for obtaining chromosomal integration of recombinant constructs have been described, for example, by Gerhardt *et al.*, METHODS FOR GENERAL AND MOLECULAR MICROBIOLOGY, American Society for Microbiology, Washington, D.C., 1994; Link *et al.* (1997) *J. Bacteriol.* **179**:6228-6237; and Metcalf *et al.* (1996) *Plasmid* **35**:1-13.

A coding sequence, as present in a recombinant construct, can encode a full-length gene product (*i.e.*, the length normally found in the wild-type cell) or any fragment of a gene product. A gene product can be a RNA or a polypeptide; untranslated RNA gene products can include structural, catalytic and regulatory RNA molecules. Examples of untranslated RNA gene products include, but are not limited to, tRNA, rRNA, antisense RNAs and ribozymes. In one embodiment, a coding sequence comprises a gene, which can encode a virulence factor, a resistance factor, or a gene product whose function is essential for a cell under a particular set of environmental conditions. Any gene of interest can be placed in operative linkage with *raf* regulatory region sequences, so that its expression is regulated by the *raf* regulatory region sequences.

In one embodiment, the invention provides recombinant constructs capable of regulating the expression of coding sequences in a host cell. These constructs comprise one or more *raf* regulatory sequences in operative linkage with a coding sequence. The constructs are suitable for use in any cell in which *raf* operon regulatory sequences are functional. Since the *raf* regulatory proteins RafR and RafS can be introduced into a cell along with, or as part of the above-mentioned recombinant construct, regulation of a coding sequence by a *raf* regulatory sequence will be attainable in many cells, which can include both Gram-positive and Gram-negative microorganisms. In preferred embodiments, the host cell is a Gram-positive microorganism, such as, for example, *Streptococci, Staphylococci, Enterococci, and Lactococci.* Preferred species include, for example, *S. pneumoniae, S. pyogenes, S. agalactiae*, Lancefield group A streptococci, Lancefield group B streptococci, Lancefield group C streptococci, Lancefield group F streptococci, Lancefield group G streptococci, and viridans streptococci. Preferred non-streptococcal species in which the *raf* regulatory system can be utilized for regulated expression of coding sequences include, for example, enterococci such as *E. faecalis,* and *E. faecium*, and lactococci such as *L. lactis*.

In the practice of one aspect of the invention, the recombinant construct is introduced into a host cell to provide regulated expression of a coding sequence.

Introduction of the construct into a host cell is performed by methods that are well-known to those of skill in the art, including, for example, natural or artificial transformation, transduction, conjugation, microinjection, transfection, electroporation, $CaPO_4$ co-precipitation, DEAE-dextran, lipid-mediated transfer, particle bombardment, *etc.*

Host cells are cultured in any suitable growth medium, including liquid or solid media. Appropriate growth media for various types of microorganisms are well-know to those of skill in the art. *See*, for example, Bergey's Manual of Systematic Bacteriology, vol. 2, Williams & Wilkins, Baltimore, 1980; Gerhardt *et al.* "Methods for General and Molecular Microbiology," American Society for Microbiology, Washington, D.C., 1994; and Murray *et al.*, *supra*.

Modulator substances can be added to the growth medium to influence the transcriptional activity of *raf* regulatory sequences. Such effects will be manifested as changes in the expression level of a coding sequence to which the *raf* regulatory sequences are operatively linked. The modulator substances can be generally characterized as inducers, which increase transcriptional activity, or negative modulators, which decrease transcription. In one embodiment of the invention, a modulator substance is a metabolite; in a preferred embodiment, it is a carbon source, in a more preferred embodiment, it is a sugar and, in a particularly preferred embodiment, raffinose serves as an inducer and maltose as a negative modulator.

The present invention utilizes systems which provide low basal expression levels and a high degree of induction. Such methods and compositions can be used, for example, to identify compounds which inhibit the growth of a microorganism, and for discovery of drug targets, including genes involved in virulence and drug resistance. Because the *S. pneumoniae raf* regulatory system is characterized by an induction level of at least 1,000-fold over a low basal expression level (see Example 14), it is well-suited for use in the practice of the invention.

Accordingly, *raf* regulatory sequences can be used to identify an essential gene of a microorganism, to regulate the level of expression of an essential gene,

and to identify inhibitors of essential genes and gene products, as disclosed herein. These features are accomplished by fusing *raf* regulatory sequences, such as P_{rafR}, P_{rafE}, P_{AGA} or others, to a homologous or heterologous coding sequence encoding an essential gene, such that the coding sequence is under the transcriptional control of the *raf* regulatory sequences. Essential genes can include those which are essential for the growth of *S. pneumoniae*, or those which are essential for the growth of any other microorganism. The essentiality of a gene may depend on the *in vivo* or *in vitro* environment of the cell in which it is expressed. For example, in the presence of an effective concentration of an antibiotic, a gene encoding resistance to that antibiotic is an essential gene. Example 7 shows the construction and properties of a fusion between the *S. pneumoniae* α-galactosidase promoter (P_{AGA}) and the *S. pneumoniae* leader peptidase (*spi*) gene. The fusion places the *spi* gene under raffinose control, limiting cell growth at low raffinose concentrations.

Additional positively-regulated promoter/activator systems will also find use in the practice of the invention. These include, but are not limited to, those for rhamnose utilization (regulated by the *rha*S or *rha*R gene products), melibiose utilization (regulated by the *mel*R gene product), xylose utilization (regulated by the *xyl*R gene product), *p*-hydroxyphenylacetic acid utilization (regulated by the *hpa*A gene product), and urease production (regulated by the *ure*R gene product). *See* Gallegos *et al.*, *supra*, for additional examples of positively regulated systems. Additional regulatable promoters that will be of use in the present invention will be well-known to those of skill in the art. They include, but are not limited to, *lac*, which is regulated by lactose and glucose; *trp*, which is regulated by tryptophan, *tac*, which is regulated by lactose, *tet*, which is regulated by tetracycline and tetracycline analogues, *gal*, which is regulated by galactose, T7, which is regulated by provision of T7 RNA polymerase, T3, which is regulated by provision of T3 RNA polymerase, SP6, which is regulated by provision of SP6 RNA polymerase, λp_R, which is regulated by λ repressor (the cI gene product), and λp_L, which is regulated by λ repressor (the cI gene product). Additional promoters from Gram-negative organisms which can be tested for their degree of

regulatability and can be useful in the practice of the invention include, but are not limited to, *lpp, phoA, recA, proU, cst-1, tetA, cadA, nar, lpp-lac, cspA,* T7-*lac,* pL-T7, T3-*lac*, T5-*lac, nprM-lac,* VHb, promoters regulated by two-component regulatory systems, and promoters regulated by the *araC/XylS* family of regulators. Two-component systems include those which utilize protein phosphorylation as a mechanism of signal transduction. In one embodiment, a sensor protein is phosphorylated upon receipt, by the cell, of an environmental stimulus. The phosphate group is then transferred to a regulator protein that undergoes a phosphorylation-induced conformational change which elicits a response such as, for example, gene transcription. *See,* for example, Makrides, *supra*; J.A. Hoch and T.J. Silhavy, *eds.* (1995) "Two-Component Signal Transduction," American Society for Microbiology, Washington, D.C.; and Gallegos *et al., supra.* Additional promoters from Gram-positive organisms which can be tested for their degree of regulatability and can be useful in the practice of the invention include, but are not limited to, *spac-1, xylA, lacA, lacR,* P15, *dnaJ, sodA, prtP, prtM,* PA170, *trpE, nisA, nisF malX, malM, xyl,* and bacteriophage promoters from φr1t and φ31. *See*, for example, de Vos *et al., supra.* Although some of these promoters are not capable, using current techniques, of basal expression levels as low as those that can be obtained with P_{BAD}, they will find use in less-preferred embodiments of the invention.

Construction of gene fusions

In a preferred embodiment of the invention, fusion of a heterologous regulatory element to a gene encoding an essential cellular function is accomplished by insertion of an *ara* regulatory cassette into the chromosome of the organism under study, or insertion of an *ara* regulatory cassette into a plasmid resident in the organism under study. The *ara* regulatory cassette can include a DNA molecule containing, in the following order, the *ara*C gene, P_C (the *ara*C promoter) and P_{BAD} (the promoter regulating expression of the *ara*B, *ara*A, and *ara*D genes). This is the order in which these elements are arranged on the *E. coli* and *S. typhimurium* chromosomes, in which the P_C and P_{BAD} promoters are

adjacent and oriented divergently. Insertion of this cassette will provide AraC function to the cell and place downstream coding sequences under the control of P_{BAD}, which is regulated by AraC. Alternatively, a cassette containing only P_{BAD} can be inserted, if AraC function is already provided by the cell.

In another embodiment, a cassette containing a gene or nucleotide sequence of interest can be inserted into the chromosome adjacent to P_{BAD} such that the gene or sequence comes under the transcriptional control of P_{BAD}. *See* Example 13. Furthermore, it will be apparent to one of skill in the art that a fusion between P_{BAD} (or any other regulatory element) and a gene or nucleotide sequence of interest can itself be moved to any one of a number of different chromosomal or extrachromosomal locations, using techniques that are well-known in the art.

By way of example, fusions can be obtained by random insertion of an *ara* regulatory cassette into a chromosome or a plasmid of a microorganism, followed by screening for strains dependent on arabinose for growth. Arabinose-dependent strains will be those in which sequences encoding an essential cellular function have been fused to the *ara* regulatory cassette in such a way that the coding sequences have come under P_{BAD} control. A coding sequence can encode a full-length gene product (*i.e.*, the length normally found in the wild-type cell) or any fragment of a gene product capable of encoding an essential cellular function. A gene product can be a RNA or a polypeptide; untranslated RNA gene products can include structural, catalytic and regulatory RNA molecules.

Random chromosomal integration is typically achieved using transposons. Transposons are DNA segments which have the ability to insert randomly within the a chromosome or plasmid of a host organism. Very little homology is required between the ends of a transposon and its integration site and the process is independent of the host's homologous recombination system. Transposon insertion is typically monitored by selection for an antibiotic resistance marker carried on the transposon. Because transposons can have low site specificity, they are widely used for random inactivation by gene disruption.

Efficient targeted chromosomal integration of an exogenous sequence, involving site-specific recombination, typically requires a stretch of homology of 200 base pairs or more and utilizes the host's homologous recombination system to achieve integration. Targeted integration typically involves a recombination event between the chromosome and a conditionally replication-defective plasmid containing chromosomal sequences and an antibiotic resistance marker. Under conditions that are non-permissive for plasmid replication, and in the presence of selective agent, the majority of surviving cells are those in which targeted recombination has occurred between the homologous sequences in the plasmid and the chromosomal DNA. Gerhardt *et al.*, supra; Link *et al.* (1997) *J. Bacteriol.* 179:6228-6237; and Metcalf *et al.* (1996) *Plasmid* 35:1-13. The same considerations apply to targeted insertion within a plasmid.

By way of example, one method for generating a fusion of an *ara* regulatory cassette to a cellular coding sequence is by flanking the regulatory cassette with sequences homologous to the targeted coding sequence, as described in Example 1, *infra*; however, other methods for generating gene fusions will be known to those of skill in the art. *See*, for example, Casadaban *et al. Meth. Enzymology*, vol. 100 (ed. R. Wu, L. Grossman, K. Moldave) Academic Press, New York, 1983) pp. 293-308; Silhavy *et al., supra*; and Gerhardt *et al., supra.* Additional embodiments of the invention include extrachromosomal gene fusions residing, for example, on plasmids. Such plasmid fusions can be constructed *in vivo* or *in vitro*, using techniques of genetics and recombinant DNA which are well-known to those of skill in the art. *See*, for example, Sambrook *et al., supra*; Ausubel, *et al.*, *supra*; Silhavy *et al.*, *supra*; and Gerhardt *et al., supra.* For the purposes of the present invention, nucleic acids constructed *in vitro* can be introduced into cells by methods that are well-known in the art, including transformation with naked DNA, electroporation, microinjection, calcium phosphate-mediated transfer, DEAE-dextran-mediated transfer, gene gun, *etc.*, to generate transformed cells.

It is clear that methods similar to those described above for *ara* regulatory cassettes can also be applied to the construction and integration of regulatory

cassettes comprising *mal*, *raf*, or other regulatory elements which allow controlled expression of a gene to which they are operatively linked.

Methods for controlled, low-level expression

In a preferred embodiment of the present invention, expression of an essential gene is regulated to a low basal level. A low basal level is less than 50% of wild-type, preferably, less than 30%, more preferably, less than 20%, and, most preferably, less than 10%. In some cases, expression of an essential gene at a low basal level will render a cell non-viable; in other cases, it will render a cell hypersusceptible to a biologically-active agent. An example of, low, basal-level regulation by *ara* P_{BAD} is provided by Guzman *et al.*, *supra*. Regulation of cell growth by arabinose, in a strain containing a P_{BAD}-*murA* fusion, is demonstrated in Example 2, *infra*.

Regulation is accomplished by fusion of a target gene to a heterologous regulatory element whose expression can be exogenously controlled, for example, by environmental conditions such as chemicals, nutrients, temperature, pH, osmolarity, *etc*. In a preferred embodiment, regulation is such that the level of the target gene product is proportional to the concentration or level of the environmental agent that is used for regulation. In a still more preferred embodiment of the invention, a target gene is fused to P_{BAD} and regulation is achieved by adjusting the concentration of L-arabinose in the growth medium. Low levels of expression are correlated with low concentrations of arabinose and/or the presence of glucose in the medium.

In additional embodiments, regulation is achieved by varying the concentration of an inducer other than arabinose. For example, regulation by maltose is achieved, in cells expressing MalR function, when a target gene is fused to *mal Pm* or *mal Px*. To provide yet another example, fusion of a target gene to the *raf* regulatory element P_{AGA} allows regulation by raffinose in a cell expressing RafR function. Additional regulatory elements in the *raf* regulatory system include the promoter of the *rafR* gene, P_{rafR}, and the promoter of the *rafE* gene, P_{rafE}. On the basis of these examples, it will be clear to one of skill in the

art that any regulatable promoter, whether positively or negatively regulated, can be used to control the expression of a target gene in response to a substance or environmental condition that regulates that particular promoter.

In diploid organisms, controlled regulation of gene expression may not be easy to achieve with a single chromosomal insertion, as it is in prokaryotes. However, in certain situations, mutation or "knockout" of one of the two copies of a target gene may lower expression of the target sufficiently for newly-acquired degrees of drug sensitivity to be obtained. Alternatively, mutation or "knockout" of one copy of a target gene, coupled with controlled expression of the remaining wild-type copy, may be used to achieve heightened drug sensitivity in a diploid. Similarly, situations may be encountered, in both prokaryotes and eukaryotes, in which multiple copies of a gene are present (*e.g.*, ribosomal genes in *E. coli*). In these situations, knockout and/or inactivation of all but one copy of the gene will allow regulation of that remaining functional copy according to the methods of the invention.

Exemplary applications

The methods and compositions of the present invention allow one to control the susceptibility of a cell to a test compound by controlling the amount of a gene product (the target) that is expressed in the cell. This is achieved by adjusting the concentration of an inducer, which will, in turn, regulate the expression of a coding sequence that is fused to a heterologous regulatory element. Sensitivity to a test compound is then determined at various levels of expression of the coding sequence and at different concentrations of the test compound. Expression of lower-than-normal levels of the target will cause a cell to become hyper-susceptible to a compound which interacts with that particular target. Alternatively, a cell expressing lower-than-normal levels of a particular gene product may become susceptible to a compound to which it is not normally susceptible (*i.e.*, to which the cell is not susceptible when it is expressing normal levels of the target). Such a compound is a candidate therapeutic which, following chemical modification, may become capable of inhibiting the viability

of cells expressing normal levels of the target. Techniques for chemical modification of potential therapeutic compounds are well-known to those of skill in the art. *See*, for example, Morin *et al.*, Chemistry and Biology of beta-Lactam Antibiotics, Academic Press, New York, 1982.

Many compounds that interact with a target are not active against wild-type cells because an intracellular concentration of the compound cannot be achieved that is sufficient for inhibition, given the concentration of target in the cell. Many factors are responsible for this type of natural resistance, for example, compounds can be hydrolyzed, effluxed, absent because lack of suitable transport, *etc.* *See*, for example, Davies (1994) *Science* **264**:375-382; and Nikaido (1994) *Science* **264**:382-388. However, inactivation of genes that are involved in natural resistance, for example efflux pumps, allows the construction of mutants that are susceptible to compounds to which wild-type cells are resistant. Such mutants can become hypersusceptible to many unrelated compounds and have been used to characterize novel antimicrobial agents. The methods and compositions of the invention can be used in concert with mutants which display increased susceptibility to compounds because of a mutation in a gene involved in metabolism, transport, efflux, and the like, to identify inhibitors that would not otherwise be detectable. See Example 11.

Figure 4 shows idealized results for a situation in which a compound inhibits cell viability by interacting with a single monomeric target, and target expression level is regulated. The figure depicts the relationship between inducer concentration and the minimum inhibitory concentration (MIC) of a test compound. The MIC is determined by assessing the minimal concentration of test compound that will inhibit growth, in the presence of a specified concentration of inducer (if applicable), typically using serial two-fold dilutions of test compound. Growth can be recorded, for example, by spectrophotometry or visual inspection of cultures. The minimum amount of test compound that completely inhibits growth, or supports less than 10% growth compared to a control culture, is defined as the MIC. For a wild-type cell, the MIC is constant at all concentrations of inducing agent, because the expression level of the target is not expected to

vary with inducer concentration. For a cell in which target expression is regulated by inducer concentration, there will be a range of inducer concentrations at which MIC is directly proportional to inducer concentration. Thus, lower inducer concentrations, which result in expression of target at levels lower than wild-type, will be correlated with lower MICs than those observed with wild-type cells. Since this assay relies solely on control of target levels, it provides a screen for candidate therapeutics regardless of whether the function of the target is known. Example 3, *infra*, shows that, for cells in which MurA expression is regulated by arabinose, MIC values both below and above the MIC for wild-type cells can be obtained, when cells are challenged with fosfomycin, a MurA inhibitor.

In certain situations, a target can be part of a multimeric structure composed of different subunits (*e.g.*, a heteromultimer), and a test compound can interact with a sub-region of the multimer contributed by more than one of the subunits, one of which is the target. If target levels are regulated, there will be, as in the situation described in the previous paragraph, a certain range of inducer concentrations at which MIC is proportional to inducer concentration. At the target levels specified by this range of inducer concentrations, the target is the limiting component of the multimeric structure. However, with increasing inducer concentration, and concurrent higher target levels, a point will be reached at which the target is no longer the rate-limiting component of the multimer. At this point, the relationship between MIC and inducer concentration reaches a plateau value, which is independent of the MIC for the wild-type strain. This situation is shown schematically in Figure 5.

Prior art screening methods are not applicable to situations in which a particular compound has multiple targets within a cell, with each target having a different degree of sensitivity to the compound. In these cases, methods of the prior art would detect effects only on the target that is most easily inhibited under the assay conditions. The methods of the present invention can be used to control the expression of a target of an inhibitor. If the target is the only cellular gene product that is targeted by the inhibitor, increasing levels of expression of the target will result in higher MICs for the inhibitor (see Figure 4). If the inhibitor

has additional targets, increase of MIC as a function of inducer concentration (*i.e.*, target levels) will reach a plateau value, indicating inhibition of a second target by the inhibitor. Figure 5 shows an idealized depiction of the data that would be obtained in such a situation. Fixing expression of the first (most sensitive) gene product, while varying expression of the other gene product(s), will allow detection of additional targets.

Thus, if a test compound interacts with a single target, the relationship between MIC and inducer concentration will be proportional at all inducer concentrations that are consistent with cell growth. *See* Examples 3, 4 and 8-13, *infra*. By contrast, if a test compound interacts with multiple targets, or with a structure formed by multiple molecules, one of which is the target, the relationship between MIC and inducer concentration will reach a plateau value at inducer concentrations at and above which the target is no longer the limiting component. Targets can be polypeptides and/or nucleic acids. For example, ribosomes contain both types of target.

Cells can be exposed to any compound that is known in the art or to be synthesized, and the route of exposure can be, for example, by inclusion of the compound in a liquid cell culture medium, by incorporation of the compound into a solid culture medium, or by application of the compound to a solid culture medium, for example, by application to the medium of a porous disc that is saturated with the compound, or by simply pipetting droplets of the compound onto a solid medium.

Use of the invention will allow the rapid identification of potential new therapeutics, such as antibacterial agents. *See* Example 5, *infra*. Candidates identified by this method can be subjected to chemical modification as known in the art (*see*, for example, Bristol, J.A. (ed.) Annual Reports in Medicinal Chemistry, Academic Press, San Diego) and tested against cells expressing normal levels of the target. Modified compounds that exhibit activity against cells expressing normal target levels are candidate therapeutics.

EXAMPLES

The following examples are intended to illustrate, not to limit the invention.

Example 1: Construction of a strain containing a fusion of P_{BAD} to *murA*

The *murA* gene was selected for testing because it encodes a cytoplasmic protein that is the target of the drug fosfomycin. An *E. coli* strain carrying a single functional copy of the *murA* gene under arabinose control in the chromosome was constructed. In this strain, the level of expression of *murA* is controlled by the amount of arabinose present in the medium. In addition, this strain is dependent on arabinose because this sugar is required to induce the expression of the essential gene *murA*, and this strain cannot metabolize arabinose because the catabolic genes have been deleted *(Δ(araCBA)araD)*. The susceptibility of the strain to fosfomycin, a uridine diphospho-N-acetyl-D-glucosamine enolpyruvyl transferase (MurA) inhibitor, was tested at different concentrations of arabinose. Susceptibility to unrelated antibiotics that inhibit other targets, *i.e.*, tetracycline (a protein synthesis inhibitor) and ciprofloxacin (a DNA gyrase inhibitor), was also investigated. A new inhibitor of MurA was identified by the practice of the invention. *See* Example 5.

The *E. coli* strains used for this example are *E. coli* VECO2042 (*pir*$^+$, *recA*); *E. coli* VECO2054 (*Δ(araCBA)araD*) and *E. coli* VECO2055 ((*Δ(araCBA)araD*) *PmurA*::Km-*ara*C-P_{BAD}).

VECO2055 was constructed as follows:

Allele replacement requires a double recombination event to occur. Two regions of homology used for recombination were the *murA* coding region and 400 base pairs of DNA immediately upstream of *murA*. The chromosomal replacement cassette and the strategy used to replace wild type *murA* with P_{BAD}-*murA* is diagrammed in Figure 6.

A DNA sequence containing *murA* was PCR-amplified from *E. coli* strain JM109 chromosomal DNA using oligonucleotides DYV-055 (SEQ ID NO. 3,

Table 2) and DYV-056 (SEQ ID NO. 4, Table 2), and cloned as an NcoI/XbaI fragment into the expression vector pBAD-MycHisB (Invitrogen Corporation, Carlsbad, CA).

400 base pairs of upstream *murA* sequence was PCR amplified from *E. coli* strain JM109 chromosomal DNA using oligonucleotides DYV-057 (SEQ ID NO. 5, Table 2) and DYV-058 (SEQ ID NO. 6, Table 2), and directly cloned into pCR2.1 (Invitrogen Corporation, Carlsbad, CA).

The suicide vector pWM95 (Metcalf *et al.* (1996) *Plasmid* **35**:1-13) was chosen to perform the allele replacement procedure. pWM95 is an ampicillin-resistant, conditionally replicative plasmid requiring the *pir* gene in *trans* for plasmid replication to occur. pWM95 also carries the *sacB* gene which confers sucrose sensitivity to transformed strains grown in the presence of sucrose. When this plasmid is introduced into a host that does not supply the Pir protein, strains carrying chromosomal integrants can be selected. *The sacB* gene then allows for selecting plasmid-free segregants as sucrose-resistant clones. *E. coli* strain VECO2042 (*pir*$^+$, *recA)* was used for all cloning steps with the conditionally replicative pWM95 and its derivatives.

The *araC*-P_{BAD}-*murA* and upstream *mur* sequences were cloned into the suicide vector pWM95 by three-way ligation to create pDY-10. The kanamycin resistance gene from plasmid pBSL99 (ATCC 87141) was cloned as a HindIII fragment into pDY10 to create pDY11.

pDY11 was introduced into *E. coli* strain VECO2054. Transformants were plated on LB plates supplemented with kanamycin (25 μg/ml) and ampicillin (100 μg/ml) and incubated at 37°C overnight. A number of transformants were streaked onto LB plates supplemented with kanamycin (25 μg/ml) and arabinose (0.2%) and incubated at 37°C overnight. Isolated colonies were then streaked onto LB plates supplemented with kanamycin (25 μg/ml), sucrose (6%), and arabinose (0.2%) to select for sucrose-resistant recombinants. NaCl was omitted from LB plates during sucrose resistance selection. Sucrose-resistant recombinants were screened for ampicillin sensitivity and arabinose growth dependence. Chromosomal replacement in candidate clones was verified by

checking the chromosomal junctions with PCR primer pairs DYV-070 / DYV-073 (SEQ ID NO: 7/SEQ ID NO: 9) and DYV-082/DYV-071 (SEQ ID NO: 10/SEQ ID NO: 8).

Table 2. Oligonucleotides used for PCR

Primer	Sequence (5' -> 3')	SEQ ID NO.
DYV-055	GGCCATGGATAAATTTCGTGTTCAGG	3
DYV-056	GGTCTAGATTATTCGCCTTTCACACGC	4
DYV-057	GGACCCGGGTCTGATTTATCAGCGAGGC	5
DYV-058	GCCATATGTCCGGAAGCTTAGTTTGTTCTCAGTTAAC	6
DYV-070	CCGGATATGGCGTTAACCG	7
DYV-071	CCCATGGTTCCAGTAAGTTCC	8
DYV-073	GTGAATGATGTAGCCGTC	9
DYV-082	CTCGCTAACCAAACCGGTAACC	10

Note: Underlined sequences correspond to non-complementary bases.

Example 2: Regulation of growth of VECO2055 by arabinose

Growth of the P_{BAD}-*mur*A fusion strain (*E. coli* VECO2055) was tested as a function of arabinose concentration. Figure 7 shows that growth of the fusion strain is dependent on arabinose concentration, demonstrating the regulation of *mur*A by a heterologous regulatory element and indicating that, at low arabinose concentrations, *mur*A function is limiting for cell growth.

Example 3: Susceptibility of VECO2055 to fosfomycin

An experiment was conducted, using the P_{BAD}-*mur*A fusion strain (VECO2055) and its parent strain (VECO2054), to compare their susceptibility to fosfomycin at different arabinose concentrations. Fosfomycin is an antibiotic which targets the *mur*A gene product. Kahan *et al.* (1974) *Ann. NY Acad. Sci.* **235**:364-386.

Preparation of inoculum

Cells were grown overnight in 5 ml of LB supplemented with 0.1% arabinose on a rotary shaker at 35°C and 200 rpm. 100 µl of overnight culture was collected, centrifuged for 5 min. at room temperature at 14,000 rpm, and the

cell pellet was suspended in 1 ml of LB with no added arabinose. The cell suspension was diluted 1:1000 in LB and used as inoculum.

Preparation of 96 well plates for checkerboard assay

Checkerboard assays were performed in 96-well microtiter plates, in which arabinose concentration was varied in the first dimension and fosfomycin concentration was varied in the second dimension. Fosfomycin concentrations varied by two-fold between rows; dilutions were performed in LB supplemented with different concentrations of arabinose, or lacking arabinose. A control row lacking fosfomycin, and a control column lacking arabinose, were also included. Total volume in each well was 50 µl.

Inoculation of plates and incubation

50 µl of inoculum, (*i.e.*, diluted cell suspension) was added to each well. After 20 hours of incubation at 35°C, cell growth was measured in each well and compared to wells with no inoculum.

The results, presented in Figure 8, show the minimum inhibitory concentration (MIC) of fosfomycin, as a function of arabinose concentration, for the P_{BAD}-*mur*A fusion strain *E. coli* VECO2055, compared to wild-type. The results demonstrate that MIC values above and below the MIC for wild-type cells can be attained in the fusion strain through adjustment of arabinose levels.

Example 4: Comparison of susceptibility of VECO2055 to fosfomycin with susceptibility to antibiotics which do not target the *mur*A gene product

An experiment was conducted on the P_{BAD}-*mur*A fusion strain (*E. coli* VECO2055) to compare its susceptibility to fosfomycin with its susceptibility to several other antibiotics (tetracycline and ciprofloxacin) which do not target the *mur*A gene product.

Preparation of inoculum

Cell culture and preparation of inocula were performed as described in Example 3, *supra*.

Preparation of 96 well plates for checkerboard assay

Checkerboard assays were performed in 96-well microtiter plates, in which arabinose concentration was varied in the first dimension and antibiotic concentration was varied in the second dimension. Antibiotic concentrations varied by two-fold between rows; dilutions were performed in LB supplemented with different concentrations of arabinose or lacking arabinose. A control row lacking antibiotic, and a control column lacking arabinose, were also included. Total volume in each well was 50 µl. Similar assays were conducted, using fosfomycin, tetracycline or ciprofloxacin, to test the influence of arabinose on susceptibility to these antibiotics.

Inoculation of plates and incubation

50 µl of inoculum, (*i.e.*, diluted cell suspension) was added to each well. After 20 hours of incubation at 35°C, cell growth was measured in each well and compared to wells with no inoculum. Figure 9 shows that the susceptibility of strain VECO2055 to tetracycline and ciprofloxacin was independent of the presence of arabinose in the medium. The variation observed for sensitivity to ciprofloxacin and tetracycline is typical of that obtained in a MIC determination. Sensitivity to fosfomycin was dependent on arabinose concentration, confirming the results shown in Figure 8. The minimum concentration of arabinose needed to support growth was 5×10^{-5}%.

The susceptibility of strain VECO2055 to fosfomycin was strongly associated with the amount of arabinose present in the medium. MICs comparable to the wild type were achievable when sufficient arabinose was added to the medium. Cells became more susceptible to fosfomycin at lower concentrations, in a concentration-dependent manner. The difference in susceptibility between the lowest and highest MICs was approximately 100-fold. The susceptibility to tetracycline and ciprofloxacin remained constant and was independent of arabinose concentration.

Example 5: Identification and validation of a new inhibitor of MurA

Preparation of inoculum

Cell culture and preparation of inocula were performed as described in Example 3, *supra*. *E. coli* VECO2055 was used in all experiments.

Screening

Microtiter plates were used for screening. Wells contained LB with 0, 0.0004, or 0.002% arabinose. Eighty unrelated compounds were tested; test compounds were added to the wells at final concentrations of 2, 4 and 8 μg/ml. 50 μl of inoculum was added to each well and the plates were incubated at 35°C. Growth was measured after 24 and 40 hours. The growth of cells incubated at the higher arabinose concentration (0.002%) was not inhibited by any of the test compounds. When cells were incubated at the lower arabinose concentration (0.0004%), compound 47-7-70 was the only test compound that inhibited growth at all three concentrations tested. As a control, the MIC for fosfomycin was determined at each arabinose concentration.

Inhibition of MurA activity

The enzymatic activity of purified MurA was assayed in the presence of the same test compounds used in the screen described *supra*. Different concentrations of the test compounds were added to buffer (50 mM Tris-HCl, pH 8; 0.2 mM UDP-N-acetylglucosamine) containing 7 μg/ml MurA. The reaction was started by addition of phosphoenolpyruvate to 0.1 mM, and the reaction mixture was incubated for 30 min at 25°C. Released phosphate was measured with malachite green reagent, and quantitated by spectrophotometry. Only compound 47-7-70 showed inhibitory activity, with an IC_{50} of 8 μg/ml.

Susceptibility of VECO2055 to compound 47-7-70 as a function of arabinose concentration

A checkerboard assay was performed in a 96-well microtiter plate, similar to that described in Example 3, *supra*. Concentration of compound 47-7-70 was varied in one dimension, and arabinose concentration was varied in the other dimension. A control row lacking compound 47-7-70, and a control column

lacking arabinose, were also included on the plate. In addition, a control plate containing dilutions of fosfomycin instead of compound 47-7-70 was also tested.

50 µl of VECO2055 inoculum, prepared as described in Example 3, *supra*, was added to each well. The plates were incubated 20 hours at 35°C, at which time cell growth was measured and compared to that in wells that had not received an inoculum. The results are presented in Table 3. As can be seen, increasing susceptibility of VECO2055 to compound 47-7-70 was correlated with lower concentrations of arabinose in the medium, as expected for a compound that blocks cell growth by inhibition of the MurA enzyme.

Table 3: MIC of compound 47-7-70 as a function of arabinose concentration for VECO2055

% arabinose in medium	MIC of 47-7-70 (µg/ml)
0.00018	2
0.000375	4
0.00075	8
0.0015	16
0.00312	16

Example 6: Construction and properties of a tightly regulated maltose regulatory system

MalR is a repressor that controls the expression of the maltosaccharide regulon in *S. pneumoniae* and belongs to the LacI-GalR family of repressors. Two operons are regulated in opposite direction, *malXCD* (*Px* promoter) and *malMP* (*Pm* promoter), see Figure 1 (SEQ ID NO: 1). Affinity of the MalR for *Pm* is higher than for *Px* and, in both cases, a high basal level of expression has been reported. Nieto *et al.* (1997) *J. Biol. Chem.* **272**:30860-30865. This example shows that tighter regulation of the *Px* promoter can be obtained by modifying the repressor site and by growing cells in minimal medium.

The *S. pneumoniae* strains used for this example are *S. pneumoniae* VSPN3026, *S. pneumoniae* VSPN3021 with *katA* under *Px* control, *S.*

pneumoniae VSPN3025 with *katA* under modified *Px* control, and *S. pneumoniae* VSPN3022 with *katA* under *Px* control, with *Pm* upstream.

Construction of VSPN3021

Px was PCR amplified from VSPN3026 using oligonucleotides MAL1 and MAL2 (Table 4), and cloned into the T-tailed PinPoint Xa-1 T-Vector (Promega Corporation, Madison, WI). The construct was digested with EcoRI and BamHI, and the 387 bp *Px*-containing fragment was cloned into pR326 vector to create pR326MX. The reporter gene used for measuring expression was the gene for catalase, *katA,* from *B. subtilis* ATCC 6633. The gene encoding *katA* was PCR-amplified using oligonucleotides KAT1 and KAT2 (Table 4) and cloned into the NdeI site of pR326MX to create pR326MXK.

An additional DNA sequence was added to the construct to target the insertion into a non-essential DNA sequence of the *S. pneumoniae* chromosome. For this purpose, a 300 bp fragment of the *cpbA* gene was PCR-amplified from the DNA of *S. pneumonia* VSPN3026, using oligonucleotides CPB1 and CPB2 (Table 4), and the amplification product was inserted into the ClaI site of pR326MXK to create pR326MXKC. This plasmid was used to transform VSPN3026 and construct a *S. pneumoniae* strain, VSPN3021, that carries the insertion in the chromosome, according to the insertional duplication mutagenesis method of Claverys *et al.* (1995) *Gene* 164:123-128.

Construction of VSPN3025

A DNA sequence containing the repressor binding site of *Px* was mutagenized, to convert a GGA to a GCG (see Figure 1), by using the QuickChange site directed mutagenesis kit (Stratagene, La Jolla, CA). The oligonucleotides MAL6 and MAL6C (Table 4) were used as primers, and pR326MXKC was used as template. Note that MAL6C includes the mutant sequence. The resultant plasmid, with a mutation in *Px*, was called pR326MMXKC. This plasmid was used to transform VSPN3026 and construct a *S. pneumoniae* strain, VSPN3025, that carries the insertion in the chromosome.

Construction of VSPN3022

The *Px* and *Pm* regulatory region was PCR amplified from VSPN3026 using oligonucleotides MAL3 and MAL2 (Table 4), and cloned into the T-tailed PinPoint Xa-1 T-Vector (Promega Corporation, Madison, WI). The construct was digested with EcoRI and BamHI, and the *Px* and *Pm*-containing fragment was cloned into the pR326 vector to create pR326MXM. The catalase gene was inserted under the control of this regulatory region as described above to create pR326MXMKC. This plasmid was used to transform VSPN3026 and construct a *S. pneumoniae* strain, VSPN3022, that carries the insertion in the chromosome.

Table 4. Oligonucleotides used in *mal* constructions

Oligo	Sequence (5' -> 3')	SEQ ID NO
MAL1	TAGGTTGAATTCATAGAAAATAGATAGGGATTAGAACCA GGG	11
MAL2	TGCGAGGATCCTACTTGTCGTCGTCGTCCTTGTAGTCGAT ATCATATGTATTCCTCCCAAAGAATAGCAAGT	12
KAT1	CCATCGCATATGAGTTCAAATAAACTGACAAC	13
KAT2	CACGACATATGAATCTTTTTTAATCGGCAATCC	14
CBP1	CTGAATCGATGCAGCCACTTCTTCTAATATGGC	15
CBP2	AGCTATCGATTTTCTAACCTTGTAGCCTCAGC	16
MAL3	TAGGTTGAATTCTCGTGTGTTAAAATAATG	17
MAL6	CGCAAACGTTTGCGTTTATGAGCTTAG	18
MAL6C	CTAAGCTCATAAACGCAAACGTTTGCG	19

Catalase assay

Catalase activity was measured in cell cultures grown in medium C+Y without glucose (Tomasz (1970) *J Bacteriol.* **101**:860-871.) and minimal medium CDEN without glucose. Rane *et al.* (1940) *J. Bacteriol.* 40:695-704. Cells were collected in mid-logarithmic phase, centrifuged, and resuspended in 20 mM Tris-HCl buffer, pH 8, containing 0.25% Triton-X100. After autolysis of cells, 10 µl of extract were added to 1 ml of 1.5 mM H_2O_2, and the reaction was followed fluorometrically with scopoletin. One unit of catalase activity is 1 µmol H_2O_2 hydrolysis per min. at 22°C (see Table 5).

Table 5. Catalase activity in *mal* insertion strains

	Growth conditions			
Strain	C+Y (2% maltose)	C+Y (No maltose)	CDEN (2% maltose)	CDEN (No maltose)
VSPN3022	43	14	69	1.12
VSPN3021	60	16	62	13
VSPN3025	65	6	66	0.35

The data (expressed as units of catalase activity) show that tighter regulation of the maltose regulatory system can be obtained when using minimal medium (CDEN), and that the tightest regulation is obtained when using the modified *Px* promoter (VSPN3025).

Example 7: Raffinose-regulated expression of the *S. pneumoniae spi* gene by the *S. pneumoniae aga* promoter

A. Construction of transcriptional fusions

A DNA sequence containing *S. pneumoniae rafR* and P_{AGA} (*aga* promoter) was PCR-amplified from *S. pneumoniae* VSPN3026 chromosomal DNA using oligonucleotides REGAGAEI5' (SEQ ID NO. 20, Table 6) and REGAGANB3' (SEQ ID NO. 21, Table 6) and cloned as an EcoRI/NdeI fragment into the integration vector pR326 (Claverys, *et al.*, *supra*) to generate plasmid pR326RafRPaga. A DNA fragment containing the first 270 bp of the leader peptidase gene (*spi*) from *S. pneumoniae* R6 chromosomal DNA was PCR-amplified using oligonucleotides MALSPI5' (SEQ ID NO. 22, Table 6) and MALSPI3' (SEQ ID NO. 23, Table 6) and cloned as an NdeI/BamHI fragment into plasmid pR326RafRPaga, resulting in plasmid pR326RPASPI.

Using this plasmid as a template, a DNA fragment containing only P_{AGA} and *spi* sequences was amplified using oligonucleotides Paga5' EI (SEQ ID NO. 24, Table 6) and MALSPI3' (SEQ ID NO. 23, Table 6) and cloned into the T-tailed pGEM-T Easy Vector (Promega Corporation, Madison, WI). The construct was digested with EcoRI and the P_{AGA}-*spi* containing fragment was cloned into

the integration vector pR326 to create PR326PagaSpi. This plasmid was used to transform VSPN3026, that had been grown in C+Y lacking sucrose and supplemented with 0.2% raffinose (C+Y+Raf). Transformants were plated on TSA-sheep blood plates supplemented with chloramphenicol (2.5 μg/ml) and raffinose (0.2%) and incubated at 37°C/5% CO_2 overnight. Chloramphenicol-resistant strains were cultured in C+Y+Raf medium.

The site of insertion was verified in one of the isolates, *S. pneumoniae* strain VSPN3041. Insertion in the targeted site in the *spi* gene was verified by PCR using the primers Paga100 (SEQ ID NO. 25, Table 6) and Spi3'SPn (SEQ ID NO. 26, Table 6). This analysis indicated that VSPN 3041 carried a truncated 270 bp *spi* gene under natural promoter control and a complete *spi* gene under P_{AGA} control.

Table 6: Sequences of oligonucleotides used for construction of a P_{AGA}-*spi* fusion and targeting of the fusion to the chromosomal *spi* gene

Oligonucleotide	Sequence (5' → 3')	SEQ ID NO.
REGAGAEI5'	CCCGGAATTCAGCTTGGTAGGATTTCATAA TGTTGCC	20
REGAGANB3'	GCCGCGGATCCGCGCATATGCATTTACTTC ACCTCATCACTTTATTG	21
MALSPI5'	GGGGAATTCCATATGAATTTATTTAAAAAT TTCTTAAAAGAGTGGG	22
MALSPI3'	GCGCTCTAGATCATTTTCGTAACGAATGGT GTCG	23
Paga5'EI	GCGCCGGAATTCCATGTGCTACCTCCTACCT AACATTTTACC	24
Paga100	CTCCTACCTAACATTTTACCAT	25
Spi3'Spn	TTAAAATGTTCCGATACGGGTGATTGG	26

B. Regulation of growth of VSPN3041 by raffinose

Since VSPN3041 carries an essential gene (*spi*) under the control of P_{AGA}, the strain should be dependent on raffinose for growth. Accordingly, the effect of

raffinose on the growth of VSPN3041 and the parent isogenic strain VSPN3026 was compared.

Preparation of 96 well plates: 96-well microtiter plates were used for the experiments. Serial twofold dilutions of raffinose (in C+Y medium) were performed across columns, including one column in which no raffinose was added. Total volume in each well was 50µl.

Preparation of inoculum: VSPN3041 and VSPN3026 were grown for 4 hours at 37° C in 5 ml C+Y+raf. Ten microliters of the culture was added to 10 ml C+Y containing 1% sucrose, incubated for 6 hours at 37° C, then frozen as a 15% glycerol solution. A 1:10 dilution of the frozen stock in C+Y lacking sucrose, containing 2 x 10^6 colony forming units/ml, was used as inoculum.

Growth of bacteria in 96-well plates: 50µl of inoculum was added to each well, to give a final volume of 100 µl. Plates were incubated ate 35° C and growth was monitored every hour up to 10 hours.

Results: Figure 10 shows the growth of VSPN3041 at different raffinose concentrations, measured by optical density. It is clear that growth of the fusion strain is dependent on raffinose concentration, demonstrating that an essential gene is regulated by raffinose in VSPN3041. In the experiment shown in Figure 11, the growth of VSPN3041 on raffinose is compared to growth on sucrose (each sugar present in medium at 0.2% w/v). The results of this experiment indicate that VSPN3041 does not grow in the absence of raffinose, again demonstrating that an essential gene is positively regulated by raffinose in this strain. Figure 12 compares the growth of VSPN3041 and the parent isogenic strain VSPN3026, at different raffinose concentrations. The results indicate that VSPN3041 is raffinose dependent, while the growth of the parent strain is not dependent on raffinose.

The raffinose-dependent phenotype of VSPN3041, compared to its parent strain, indicates that P_{AGA} controls the expression of an essential gene in VSPN3041. Given that the difference between VSPN3041 and its parent is an insertion that places the *spi* gene under P_{AGA} control, the essential raffinose-regulated gene in VSPN3041 is the *spi* gene (or a gene downstream of *spi*).

Hence, the *spi* gene (or its downstream gene) in VSPN3041, regulated by the P_{AGA} promoter in response to raffinose, is limiting for cell growth at low raffinose concentrations. Since growth is dependent on the expression of this essential gene and the level of induction can be controlled, the growth of VSPN3041 can be controlled by the induction or repression of the P_{AGA} promoter.

Example 8: Construction and properties of a strain containing a P_{BAD}-*def* transcriptional fusion

A Construction of P_{BAD}-*def* transcriptional fusions in *E. coli*

The product of the *def* gene, the enzyme peptidyl deformylase, plays a major role in protein synthesis in bacteria. A DNA sequence containing the full-length *def* gene was PCR-amplified from chromosomal DNA of *E. coli* strain JM109 using oligonucleotides DYV-157 (SEQ ID NO: 27) and DYV-158 (SEQ ID NO: 28), and cloned as a NcoI/BglII fragment into expression vector pBAD-MycHisB (Invitrogen Corporation, Carlsbad, CA), to create pDY8. Oligonucleotide sequences are given in Table 7.

pDY20 was created by PCR-amplification of the kanamycin resistance cassette from plasmid pBSL99 with the primers DYV-087 (SEQ ID NO: 35) and DYV-088 (SEQ ID NO: 36) and cloned into pBlueScriptSKII$^-$ (Stratagene, La Jolla, CA) as an Xba/SacI fragment. 600 base pairs of upstream *def* sequence were PCR amplified from *E. coli* strain JM109 chromosomal DNA using oligonucleotides DYV-155 (SEQ ID NO: 29) and DYV-156 (SEQ ID NO: 30), and cloned as a SacI/AscI fragment into vector pDY20 to create pDY9. Oligonucleotide sequences are given in Table 7.

The suicide vector pKO3 (Link *et al., supra*) was chosen to perform the allele replacement procedure with the *def* gene. pKO3 is a chloramphenicol-resistant vector containing the temperature-sensitive pSC101 origin of replication and the *sacB* gene for counter-selection. pKO3-derived plasmids are incapable of autonomous replication at 43°C. When a host strain harboring a pKO3 construct is plated at 43°C on media containing chloramphenicol, chromosomal integrants can be selected. Integration of a pKO3 construct into a host chromosome at

elevated temperature occurs via homologous recombination between *E. coli* DNA cloned into pKO3 and the *E. coli* chromosome.

The *araC*-P_{BAD}-*def* cassette was excised as an NdeI/BglII fragment from pDY8, and the upstream *def*-Kanamycin cassette was excised as an Ecl136 II/NdeI fragment from pDY9. The purified fragments were cloned in a three-way ligation with SmaI/BamHI digested pKO3 to create pDY15.

pDY15 was introduced into *E. coli* strain VECO2054. Transformants were selected on LB plates supplemented with chloramphenicol (25 μg/ml) and kanamycin (25 μg/ml), incubated at 30°C overnight. A number of transformants were streaked onto LB plates supplemented with chloramphenicol (25 μg/ml), kanamycin (25 μg/ml) and arabinose (0.2%) and incubated at 43°C overnight. Isolated colonies were then streaked onto LB plates supplemented with kanamycin (25 μg/ml) and arabinose (0.2%) and incubated at 37°C overnight. Isolated colonies were then streaked onto the same medium (LB+kan+ara) supplemented with 6% sucrose and incubated at 37°C overnight to select for sucrose-resistant recombinants. NaCl was omitted from LB plates during this sucrose resistance selection step. Sucrose-resistant recombinants were screened for chloramphenicol sensitivity and arabinose-dependent growth. Chromosomal replacement of the *def* gene in a clone, VECO2065 (*araC*-P_{BAD}-*def*), was verified by assaying for specific PCR products, derived from the chromosomal junctions, with PCR primer pairs DYV-069 (SEQ ID NO: 37)/DYV-082 (SEQ ID NO: 10) and DYV-073 (SEQ ID NO: 9)/DYV-155 (SEQ ID NO: 29). See Tables 2 and 7 for primer sequences.

B. Susceptibility of VECO2065 (P_{BAD}-*def*) strain to VRC483 and other antimicrobial agents

An experiment was conducted using the VECO2065 strain and its parent strain (VECO2054), to compare their susceptibility to VRC483 over a range of arabinose concentrations. VRC483 is a compound with antibacterial activity that targets the *def* gene product. This compound was identified in a deformylase screen at Versicor and the IC_{50} is 11 nM for *E. coli* deformylase. Deformylase activity was measured as described. Rajagopalan *et al* (1997) *Biochemistry*

36:13910-13918. The susceptibility of VECO2065 to the unrelated antibiotics fosfomycin and ciprofloxacin was also tested.

1. Preparation of inoculum. Cells were grown overnight in 5 ml of LB supplemented with 0.1 % arabinose, on a rotary shaker at 35°C and 200 rpm. 100 µl of overnight culture was collected, centrifuged at 14,000 rpm for 5 min at room temperature, and the cell pellet was suspended in 1 ml of medium with no added arabinose. The cell suspension was diluted 1:1000 in medium and used as inoculum.

2. Preparation of 96 well plates. Checkerboard assays were performed in 96-well microtiter plates, in which arabinose concentration was varied in the first dimension, and the concentration of antimicrobial compound in the second. The concentration of antimicrobial varied by two-fold between rows. Dilutions were performed in medium supplemented with different concentrations of arabinose, or medium lacking arabinose. A control row lacking antimicrobial, and a control column lacking arabinose, were also included. Total volume in each well was 50 µl.

3. Inoculation of plates and incubation. 50 µl of inoculum (*i.e.*, diluted cell suspension) was added to each well. After 20 hours of incubation at 35°C, cell growth was measured in each well and compared to wells with no inoculum.

4. Results. Figure 13 shows the minimum inhibitory concentration (MIC) of VRC483, as a function of arabinose concentration, for the P_{BAD}-*def* strain VECO2065. The parent wild-type strain (VECO2054) was not susceptible to VRC483 in the range tested. VECO2065 was susceptible to VRC483 at low arabinose concentrations, and the susceptibility was inversely related to the inducer concentration. Susceptibility of VECO2065 to compounds that do not target the product of the *def* gene, such as fosfomycin and ciprofloxacin, did not change with arabinose concentration.

Example 9: Construction and properties of a strain containing a P_{BAD}-*folA* transcriptional fusion

A Construction of P_{BAD}-*folA* transcriptional fusions in *E. coli*

The product of the *folA* gene is a dihydrofolate reductase. This enzyme is involved in folate synthesis in bacteria. A DNA sequence containing the full-length *folA* gene was PCR-amplified from the chromosomal DNA of *E. coli* strain JM109, using oligonucleotides DYV-095 (SEQ ID NO: 31) and DYV-096 (SEQ ID NO: 32), and cloned as an NcoI/BglII fragment into expression vector pBAD-MycHisB (Invitrogen Corporation, Carlsbad, CA), to create pDY5. Oligonucleotide sequences are given in Table 7.

Six hundred base pairs of upstream *folA* sequence were PCR amplified from *E. coli* strain JM109 chromosomal DNA, using oligonucleotides DYV-093 (SEQ ID NO: 33) and DYV-094 (SEQ ID NO: 34), and cloned as a SacI/AscI fragment into vector pDY20 (see Example 8) to create pDY6. Oligonucleotide sequences are given in Table 7.

The *araC*-P_{BAD}-*folA* cassette was excised as a NdeI/BglII fragment from pDY5, and the upstream *folA*-kanamycin cassette was excised as an Ecl136 II/NdeI fragment from pDY6. The purified fragments were cloned in a three-way ligation with SmaI/BamHI digested pWM95 to create pDY42.

pDY42 was introduced into *E. coli* strain VECO2054. Transformants were selected on LB plates supplemented with ampicillin (100 μg/ml) and kanamycin (25 μg/ml), incubated at 37°C overnight. A number of transformants were streaked onto LB plates supplemented with kanamycin (25 μg/ml) and arabinose (0.2%) and incubated at 37°C overnight. Isolated colonies were picked and re-streaked onto LB plates supplemented with kanamycin (25 μg/ml) and arabinose (0.2%) and incubated at 37°C overnight. Isolated colonies were then streaked onto the same medium (LB+kan+ara) supplemented with 6% sucrose and incubated at 37°C for 24 hours. The plates were then incubated for an additional 24 hours at room temperature to select for sucrose-resistant recombinants. NaCl was omitted from LB plates during this selection for sucrose resistance. Sucrose-resistant recombinants were screened for ampicillin sensitivity and arabinose-

dependent growth. Chromosomal replacement of the *folA* gene in a clone, VECO2079 (*araC*-P_{BAD}-*folA*), was verified by assaying for specific PCR products, derived from the chromosomal junctions, with PCR primer pairs DYV-093 (SEQ ID NO: 33)/DYV-163 (SEQ ID NO: 38) and DYV-107 (SEQ ID NO: 39)/DYV-218 (SEQ ID NO: 40). See Table 7 for primer sequences.

B. Susceptibility of VECO2079 (P_{BAD}-*folA*) strain to trimethoprim and other antimicrobial agents

An experiment was conducted using the VECO2079 strain and its parent strain (VECO2054), to compare their susceptibility to trimethoprim over a range of arabinose concentrations. trimethoprim is a compound with antibacterial activity that targets dihydrofolate reductase, the product of the *folA* gene. Huovinen *et al.* (1995) *Antimicrob. Agents Chemother.* **39(2)**:279-289. The susceptibility of VECO2079 to the unrelated antibiotics fosfomycin and ciprofloxacin was also tested.

1. Preparation of inoculum. Cells were grown overnight in 5 ml of LB supplemented with 0.1 % arabinose, on a rotary shaker at 35°C and 200 rpm. 100 µl of overnight culture was collected, centrifuged at 14,000 rpm for 5 min at room temperature, and the cell pellet was suspended in 1 ml of medium with no added arabinose. The cell suspension was diluted 1:1000 in medium and used as inoculum.

2. Preparation of 96 well plates. Checkerboard assays were performed in 96-well microtiter plates, in which arabinose concentration was varied in the first dimension, and the concentration of antimicrobial compound in the second. The concentration of antimicrobial varied by two-fold between rows. Dilutions were performed in medium supplemented with different concentrations of arabinose, or medium lacking arabinose. A control row lacking antimicrobial, and a control column lacking arabinose, were also included. Total volume in each well was 50 µl.

3. Inoculation of plates and incubation. 50 µl of inoculum (*i.e.*, diluted cell suspension) was added to each well. After 20 hours of incubation at 35°C, cell growth was measured in each well and compared to wells with no inoculum.

4. Results. Figure 14 shows the minimum inhibitory concentration (MIC) of trimethoprim, as a function of arabinose concentration, for the P_{BAD}-*folA* strain (VECO2079) compared to the parent wild-type strain, VECO2054. The results show that the MIC of trimethoprim, a *folA* inhibitor, was dependent on arabinose concentration in the P_{BAD}-*folA* strain (VECO2079); while MIC values of trimethoprim for the wild-type strain were not dependent on arabinose concentration. Figure 14 also shows that susceptibility of VECO2079 to compounds that do not target the product of the *folA* gene, such as fosfomycin and ciprofloxacin, did not change with arabinose concentration.

Example 10: Construction and properties of a strain containing a P_{BAD}-*gyrB* transcriptional fusion

A. Construction of P_{BAD}-*gyrB* transcriptional fusions in *E. coli.*

The product of the *gyrB* gene is the beta subunit of gyrase, a bacterial DNA topoisomerase. A DNA sequence containing the full-length *gyrB* gene was PCR-amplified from the chromosomal DNA of *E. coli* strain JM109, using oligonucleotides DYV-099 (SEQ ID NO: 41) and DYV-204 (SEQ ID NO: 42), and cloned as an NcoI/PstI fragment into expression vector pBAD-MycHisB (Invitrogen Corporation, Carlsbad, CA) to create pDY34. Oligonucleotide sequences are given in Table 7.

Six hundred base pairs of upstream *gyrB* sequence were PCR-amplified from *E. coli* strain JM109 chromosomal DNA using oligonucleotides DYV-097 (SEQ ID NO: 43) and DYV-098 (SEQ ID NO: 44), and cloned as a SacI/AscI fragment into vector pDY20 (see Example 8) to create pDY38. Oligonucleotide sequences are given in Table 7.

The *araC*-P_{BAD}-*gyrB* cassette was excised as an NdeI/ Xba fragment from pDY34, and the upstream *gyrB*-kanamycin cassette was excised as an Ecl136 II/ NdeI fragment from pDY38. The purified fragments were cloned in a three-way ligation with SmaI/XbaI digested pWM95 to create pDY40.

pDY40 was introduced into *E. coli* strain VECO2054. Transformants were selected on LB plates supplemented with ampicillin (100 μg/ml) and

kanamycin (25 μg/ml), incubated at 37°C overnight. A number of transformants were streaked onto LB plates supplemented with kanamycin (25 μg/ml) and arabinose (0.2%) and incubated at 37°C overnight. Isolated colonies were picked and re-streaked onto the same medium (LB+kan+ara) and incubated at 37°C overnight. Isolated colonies were then streaked onto the same medium (LB+kan+ara) supplemented with 6% sucrose and incubated at 37°C for 24 hours. The plates were then incubated for an additional 24 hours at room temperature to select for sucrose-resistant recombinants. NaCl was omitted from LB plates during this selection for sucrose resistance. Sucrose-resistant recombinants were screened for ampicillin sensitivity and arabinose-dependent growth. Chromosomal replacement of the *gyrB* gene in a clone, VECO2083 (*araC*-P_{BAD}-*gyrB*), was verified by assaying for specific PCR products, derived from the chromosomal junctions, with PCR primer pairs DYV-211 (SEQ ID NO: 45)/DYV-163 (SEQ ID NO: 38) and DYV107 (SEQ ID NO: 39)/DYV-214 (SEQ ID NO: 46). See Table 7 for primer sequences.

B. Susceptibility of VECO2083 (P_{BAD}-*gyrB*) strain to novobiocin and other antimicrobial agents

The susceptibility of VECO2083 strain to novobiocin, and to the unrelated antibiotics fosfomycin and ciprofloxacin, was tested. Novobiocin is an antibiotic of the coumarin group that inhibits gyrase by binding to the *gyrB* gene product. Maxwell (1993) *Mol Microbiol.* **9(4)**:681-686.

1. Preparation of inoculum. Cells were grown overnight in 5 ml of LB supplemented with 0.1 % arabinose, on a rotary shaker at 35°C and 200 rpm. 100 μl of overnight culture was collected, centrifuged at 14,000 rpm for 5 min at room temperature, and the cell pellet was suspended in 1 ml of medium with no added arabinose. The cell suspension was diluted 1:1000 in medium and used as inoculum.

2. Preparation of 96 well plates. Checkerboard assays were performed in 96-well microtiter plates, in which arabinose concentration was varied in the first dimension, and concentration of antimicrobial compound in the second. The concentration of antimicrobial varied by two-fold between rows. Dilutions were

Table 7: Oligonucleotides used for PCR

Oligo	Sequence (5' -> 3')	SEQ ID NO
DYV-157	GCATGCCATGGTTTCAGTTTTGCAAGTGTTAC	27
DYV-158	CGAAGATCTTTAGTTCTTATCCTTAAGC	28
DYV-155	GCGGAGCTCGCAGACTGGCAGCCAGTCG	29
DYV-156	TTGGCGCGCCTCCAGAGATGTGTTCAGG	30
DYV-095	GCATGCCATGGCAATCAGTCTGATTGCGGCGTTAGC	31
DYV-096	CGAAGATCTTTACCGCCGCTCCAGAATCTCAAAGC	32
DYV-093	GCGGAGCTCGGCGATGCCACGCGGATGG	33
DYV-094	GCTTGGCGCGCCAACGAGTCCACGCTCTCTCC	34
DYV-087	GGTATACCATATGCGAGCTCCAGGCGCGCCTGCAGGA ATTCGATATCAAGC	35
DYV-088	TGCTCTAGAGCCATATGTTCCGCTAGCTTCACGCTGCC	36
DYV-069	GCACCGGAATTCCCGGGTCAGCCAGTCTAACTGCGAA AGCG	37
DYV-163	CCTCGACGGTATCGATAAGC	38
DYV-107	TAGCGGATCCTACCTGACGC	39
DYV-218	CGGGATCCGCGAAGAGTACCAGTACACC	40
DYV-099	GCATGCCATGGCATCGAATTCTTATGACTCCTCC	41
DYV-204	GTCCGATCGTTAAATATCGATATTCGCCGC	42
DYV-097	GCGGAGCTCAGCGATTGCTCAAGCAGCG	43
DYV-098	GCTTGGCGCGCCTCTCGCTCATTTATACTTGGG	44
DYV-211	TCAGCGGCCGCCAGCGTGCAGATTGAAGATGC	45
DYV-214	TGACTCGAGCCGTGTAGTAGCTGATATCACGG	46
VCJ005	CCACCATAATTGACGAACGC	47
VCJ007	GTCTTCGGTACGGTCATGGTG	48

Example 11: Construction and properties of a hypersusceptible *E. coli* strain

This example describes the construction of a strain of *E. coli*, VECO2068, with an essential gene, *def*, under P_{BAD} control, and with a deletion in the *tolC* gene. Because the essential *def* gene is under P_{BAD} control, the susceptibility of VECO2068 to inhibitors of the *def* gene product depends on the concentration of arabinose in the growth medium. Mutants in *tolC* are hypersusceptible to many compounds, because *tolC* encodes an outer membrane protein, which can serve as a component of an efflux pump. Thus, the threshold for susceptibility to compounds which interact with the *def* gene product is lowered in a *tolC* mutant, compared to wild-type. Because of its heightened susceptibility, the *tolC*/P_{BAD}-*def* strain can be used for detecting compounds that otherwise would not have been identified as inhibitors of a strain that is wild-type for *tolC*.

A. Construction of a *tolC* deletion in an *E. coli* strain containing a P_{BAD}-*def* transcriptional fusion

The *tolC* gene was PCR amplified from *E. coli* strain VECO1004 using primers VCJ005 (SEQ ID NO: 47, Table 7) and VCJ007 (SEQ ID NO: 48, Table 7). The 2.7 kb PCR product was blunt-end cloned into pUC18 creating pCH12. A 700 bp internal deletion of *tolC* was created by digestion of pCH12 with the compatible enzymes PstI and NsiI, creating pCH13. The 2.0 kb Δ*tolC* fragment was excised from pCH13 by SmaI/EheI digestion and cloned into SmaI digested pKO3 creating pDY92.

The pDY92 plasmid was used to introduce the *tolC* deletion mutation (Δ*tolC*) into the chromosome of VECO2065 (an *E. coli* strain containing a chromosomal P_{BAD}-*def* fusion) via the selection /counter-selection procedure previously described for other suicide vector constructs. *See* Examples 8-10 Transformed cells were screened for successful integration of the Δ*tolC* mutation by plating on LB + 0.2% arabinose, and replica-plating onto MacConkey agar, which does not support growth of Δ*tolC* mutants. Confirmation of Δ*tolC* integration in MacConkey-sensitive clones was verified by PCR with oligonucleotides VCJ005 (SEQ ID NO: 47, Table 7) and VCJ007 (SEQ ID

NO: 48, Table 7). A Δ*tolC* strain containing the P_{BAD}-*def* fusion, VECO2068, was selected for susceptibility testing.

B. Susceptibility of VECO2068 to VRC483 and other antimicrobial agents

An experiment was conducted using VECO2068 (Δ*tolC*, P_{BAD}-*def*) and a parent strain (VECO2066) containing a deleted *tolC* gene and lacking the P_{BAD}-*def* fusion, to compare their susceptibility to VRC483 over a range of arabinose concentrations. VRC483 is a compound with antibacterial activity that targets the *def* gene product. This compound was identified in a deformylase screen at Versicor and the IC_{50} is 11 nM for *E. coli* deformylase. Deformylase activity was measured as described in Example 8 (Rajagopalan *et al., supra*).

1. Preparation of inoculum. Cells were grown overnight in 5 ml of LB supplemented with 0.1 % arabinose, on a rotary shaker at 35°C and 200 rpm. 100 µl of overnight culture was collected, centrifuged at 14,000 rpm for 5 min at room temperature, and the cell pellet was suspended in 1 ml of medium with no added arabinose. The cell suspension was diluted 1:1000 in medium and used as inoculum.

2. Preparation of 96 well plates. Checkerboard assays were performed in 96-well microtiter plates, in which arabinose concentration was varied in the first dimension, and concentration of antimicrobial compound in the second. The concentration of antimicrobial varied by two-fold between rows. Dilutions were performed in medium supplemented with different concentrations of arabinose, or medium lacking arabinose. A control row lacking antimicrobial, and a control column lacking arabinose, were also included. Total volume in each well was 50µl.

3. Inoculation of plates and incubation. 50µl of inoculum (*i.e.*, diluted cell suspension) was added to each well. After 24 hours of incubation at 35°C, cell growth was measured in each well and compared to wells with no inoculum.

4. Results. Figure 16 shows the minimum inhibitory concentration (MIC) of VRC483, as a function of arabinose concentration, for the *tolC*/P_{BAD}-*def* strain VECO2068. The results show that the MIC of VRC483, a deformylase inhibitor, for VECO2068 is dependent on the concentration of arabinose in the growth

medium. The parent wild type strain was not susceptible to the antibiotic in the range tested. Comparison to Figure 13 (Example 8) indicates that susceptibility to VRC483 occurs at lower arabinose concentrations, as predicted. Figure 16 also shows that the susceptibility of VECO2068 to compounds that do not target the *def* gene, such as fosfomycin and ciprofloxacin, did not vary with arabinose concentration.

Example 12: Construction and properties of a *S. pneumoniae* strain containing a P_{AGA}-*def* transcriptional fusion

A. Construction of P_{AGA}-*def* transcriptional fusions in *S. pneumoniae*

A DNA sequence containing *rafR* and P_{AGA} (*aga* promoter) was PCR-amplified from *S. pneumoniae* VSPN3026 chromosomal DNA using oligonucleotides REGAGAEI5' (SEQ ID NO: 20, Table 6) and REGAGANB3' (SEQ ID NO: 21, Table 6) and cloned as an EcoRI/NdeI fragment into the integration vector pR326 (Claverys *et al., supra*), resulting in plasmid pR326RafRPaga.

The first 317 bp of the deformylase gene (*def*) from *S. pneumoniae* VSPN3026 chromosomal DNA were PCR amplified using oligonucleotides MALDEF5' (SEQ ID NO: 49 Table 8) and MALDEF3' (SEQ ID NO: 50 Table 8) and cloned as an NdeI/BamHI fragment into plasmid pR326RafRPaga, resulting in plasmid pR326RPADEF.

Using pR326RPADEF as a template, a DNA sequence containing only P_{AGA} and the *def* fragment was amplified using oligonucleotides Paga5'EI (SEQ ID NO: 24, Table 6) and MALDEF3' (SEQ ID NO: 50, Table 8) and cloned into the T-tailed pGEM-T Easy Vector (Promega Corporation, Madison, WI). The construct was digested with EcoRI and the P_{AGA}-*def* containing fragment was cloned into the integration vector pR326 to create pR326Pagadef. This plasmid was used to transform VSPN3026, grown in C+Y without sucrose and supplemented with different raffinose concentrations (two-fold dilutions from 2% to 0.008 %). Transformants were used to inoculate tubes containing 2 ml C+Y medium supplemented with chloramphenicol (2.5 μg/ml) and different raffinose

concentrations (1% to 0.041%), and incubated at 37°C overnight. The overnight culture was used to inoculate (10 µl per well) a 96 well microtiter plate containing 200 µl C+Y with different raffinose concentrations, ranging from 1% to 0.041% (see above,) and incubated at 37°C in a CO_2 incubator with 5% CO_2. Cultures were plated on TSA sheep-blood agar plates containing chloramphenicol (2.5 µg/ml) and 0.2% raffinose. The plates were incubated overnight at 35°C and a single colony was picked and transferred to C+Y medium supplemented with chloramphenicol (2.5 µg/ml) and raffinose (0.03%). The resulting *S. pneumoniae* strain, VSPN3044, carries an insertion at the *def* locus in the chromosome, which was verified by PCR using the primers Paga100 (SEQ ID NO: 25, Table 6) and DEF3'Bam (SEQ ID NO: 51, Table 8). The strain carries a truncated 316 bp *def* gene under natural promoter control and a full-length *def* gene under P_{AGA} control.

Table 8: Oligonucleotides used in P_{AGA} construction and mutant characterization

Oligo	Sequence (5' -> 3')	SEQ ID NO
MALDEF5'	GGGGAATTCCATATGTCTGCAATAGAACGTATTAC	49
MALDEF3'	CCGCGGATCCAAATCGTAGGCTTCCTGTGG	50
DEF3'Bam	GGCGCGGATCCTTAAGCTTCGATTTCTGTAACCATAC CTG	51

B. Susceptibility of VSPN3044 to VRC483, vancomycin, and erythromycin

An experiment was conducted using the P_{AGA}-*def* strain VSPN3044 and its parent strain (VSPN3026), to compare their susceptibility to VRC483 as a function of raffinose concentration. Susceptibility of VSPN 3044 to vancomycin and erythromycin at different raffinose concentrations was also tested. VRC483 is a compound with antibacterial activity that targets the *def* gene product. This compound was identified in a deformylase screen at Versicor and the IC_{50} is 11 nM for *E. coli* deformylase. Deformylase activity was measured as described in Example 8 (Rajagopalan *et al., supra*).

normal chromosomal location of P_{BAD}. In this way, potential polar effects on genes downstream from the gene of interest are avoided.

A. Construction of an *E. coli* Δ*tolC* strain with a P_{BAD}-*lpxC* transcriptional fusion

The product of the *lpxC* gene is the enzyme UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase, which plays a major role in lipopolysaccharide synthesis in Gram-negative bacteria. A DNA sequence containing the full-length *lpxC* gene was PCR-amplified from *E. coli* strain MG1655 chromosomal DNA using oligonucleotides DYV-240 (SEQ ID NO: 56) and DYV-241 (SEQ ID NO: 57, Table 9), and cloned as an NcoI/BglII fragment into pNR41 creating pNR43. pNR41 contains two regions of homology with the *araBAD* locus on the chromosome. One region contains an optimized P_{BAD} promoter and approximately 500 bp of upstream DNA corresponding to the *araC* gene, the second region contains 600 bp of an internal fragment of the *araD* gene. The *araC*-P_{BAD}-*lpxC*-*araD* cassette was excised from pNR43 as an XmaI/SalI fragment and cloned into XmaI/SalI digested pKO3, thus creating pNR48.

pNR48 was transformed in *E. coli* strain MG1655. Transformants were selected on LB plates supplemented with chloramphenicol (25 μg/ml) at 30°C. A number of transformants were streaked onto LB plates supplemented with chloramphenicol (25 μg/ml) and incubated at 43°C overnight. Isolated colonies were then restreaked at 43°C onto LB plates supplemented with chloramphenicol (25 μg/ml). Isolated colonies were next streaked onto LB plates and incubated at 37°C overnight. Isolated colonies were then streaked onto LB plates supplemented with sucrose (6%) and incubated at 37°C overnight to select for sucrose-resistant recombinants. NaCl was omitted from LB plates during sucrose resistance selection. Sucrose-resistant recombinants were screened for the inability to ferment arabinose on MacConkey agar plates supplemented with 0.4% arabinose and scored for chloramphenicol sensitivity. Chloramphenicol-sensitive clones that were deficient in arabinose utilization were candidates for successful integration of the *lpxC* gene into the *araBAD* operon. One clone, VECO2520

(*araC*-P_{BAD}-*lpxC*), was verified by checking the chromosomal junctions with PCR primer pair DYV-246/DYV-249 (SEQ ID NOS: 58 and 59, Table 9).

Having placed the *lpxC* gene under arabinose control at the *araBAD* locus, the next step was to delete the endogenous *lpxC* gene from its normal chromosomal context. An in frame deletion of *lpxC*, which resides in a dicistronic operon with the essential gene *secA*, was made using crossover PCR. Link *et al.* (1997) *J. Bacteriol.* **179**:6228-6237. The crossover PCR reaction created a 1.2 kb product consisting of 600 bp fragments of DNA to the left and right of the sequence targeted for deletion. The four primers used for crossover PCR amplification were DYV-224 (SEQ ID NO: 52), DYV-225 (SEQ ID NO: 53), DYV-226 (SEQ ID NO: 54) and DYV-227 (SEQ ID NO: 55, Table 9). The resulting PCR product was digested with BamHI and cloned into BamHI-digested pKO3, creating pNR36.

Plasmid pNR36 was transformed in *E. coli* strain VECO2520. Transformants were selected on LB plates supplemented with 0.2% arabinose and chloramphenicol (25 μg/ml) at 30°C. A number of transformants were streaked onto LB plates supplemented with 0.2% arabinose and chloramphenicol (25 μg/ml) and incubated at 43°C overnight. Isolated colonies were then restreaked at 43°C onto LB plates supplemented with 0.2% arabinose and chloramphenicol (25 μg/ml). Isolated colonies were next streaked onto LB plates supplemented with 0.2% arabinose and incubated at 37°C overnight. Isolated colonies were then streaked onto LB plates supplemented with 0.2% arabinose and sucrose (6%) and incubated at 37°C overnight to select for sucrose-resistant recombinants. NaCl was omitted from LB plates during sucrose resistance selection. Sucrose-resistant recombinants were screened for arabinose-dependent growth and scored for chloramphenicol sensitivity. Chloramphenicol-sensitive clones that required arabinose for growth were candidates for successful deletion of the *lpxC* gene from its normal chromosomal context. One clone, VECO2522 (*araC*- P_{BAD}-*lpxC* ,Δ*lpxC*), was verified by checking the chromosomal junctions with PCR primer pairs DYV-224/DYV-227 (SEQ ID NOS: 52 and 55, Table 9).

The *tolC* gene was PCR-amplified from *E. coli* strain VECO1004 using primers VCJ005 (SEQ ID NO: 47, Table 7) and VCJ007 (SEQ ID NO: 48, Table 7). The 2.7 kb PCR product was blunt-end cloned into pUC18 creating pCH12. A 700 bp internal deletion of *tolC* was created by digestion of pCH12 with the compatible enzymes PstI and NsiI, creating pCH13. The 2.0 kb Δ*tolC* fragment was excised from pCH13 with SmaI-EheI and cloned into SmaI digested pKO3, creating pDY92.

The Δ*tolC* mutation was introduced onto VECO2522 and *E. coli* MJ1655 with pDY92 via the selection/counter-selection procedure previously described for other suicide vector constructs. Successful integration of the Δ*tolC* mutation was screened on MacConkey agar which does not support growth of Δ*tolC* mutants. Confirmation of Δ*tolC* integration in MacConkey-sensitive clones was verified by PCR with oligonucleotides VCJ005 (SEQ ID NO: 47, Table 7) and VCJ007 (SEQ ID NO: 48, Table 7). VECO2524 is an *araC*-P_{BAD}-*lpxC*, Δ*lpxC*, Δ*tolC* mutant and VECO2526 is a Δ*tolC* mutant. These strains were used for further experiments.

B. Susceptibility of VECO2524 (P_{BAD}-*lpxC*) to L159692 and other antimicrobial agents

An experiment was conducted using the VECO2524 strain and the *tolC* isogenic strain (VECO2526), to compare their susceptibility to L159692 within a range of arabinose concentrations. L159692 is an antibacterial compound that targets the *lpxC* gene product. Onishi *et al.* (1996) *Science* **274**:980-982. The susceptibility of VECO2526 to other unrelated antibiotics, linezolid and erythromycin, was also tested.

1. Preparation of inoculum. Cells were grown overnight in 5 ml of LB supplemented with 0.1 % arabinose, on a rotary shaker at 35°C and 200 rpm. 100 µl of overnight culture was collected, centrifuged for 5 min. at room temperature at 14,000 rpm, and the cell pellet was suspended in 1 ml of medium with no added arabinose. The cell suspension was diluted 1:1000 in medium and used as inoculum.

2. Preparation of 96 well plates. Checkerboard assays were performed in 96-well microtiter plates, in which arabinose concentration was varied in the first dimension, and antimicrobial compound in the second. Antimicrobial concentration varied by two-fold between rows. Dilutions were performed in medium supplemented with different concentrations of arabinose or lacking arabinose. A control row lacking antimicrobial, and a control column lacking arabinose, were also included. Total volume in each well was 50 µl.

3. Inoculation of plates and incubation. 50 µl of inoculum (*i.e.*, diluted cell suspension) was added to each well. After 20 hours of incubation at 35°C, cell growth was measured in each well and compared to wells with no inoculum.

4. Results. Figure 18 shows the minimum inhibitory concentration (MIC) of L159692, linezolid, and erythromycin, as a function of arabinose concentration, for the P_{BAD}-*lpxC* strain. The susceptibility of the isogenic *tolC* strain, VECO2526, to any of the compounds tested was not influenced by the amount of inducer in the medium.

Table 9: Oligonucleotides used for PCR in the construction of P_{BAD}-*lpxC* fusions

Oligo	Sequence	SEQ ID NO
DYV-224	TCGGATCCGGCTACGCAATGATGGGTTC	52
DYV-225	CCCATCCACTAAACTTAAACATGTCCTTTGTTTGATCATCG	53
DYV-226	TGTTTAAGTTTAGTGGATGGGTTGGCCTTCAAAGCGCCTTCA	54
DYV-227	GTGGATCCGTAATGCAAGATCTTGCGC	55
DYV-240	GGTTCCATGGCAATCAAACAAAGGACACTTAAACG	56
DYV-241	GTCAGATCTTTATGCCAGTACAGCTGAAGG	57
DYV-246	GACCCGGGTGATACCATTCGCGAGCC	58
DYV-249	GAGTCGACGCAGCGTTTGCTGCATATCC	59

Example 14: Regulatory properties of the *S. pneumoniae raf* gene cluster

The regulation of the *S. pneumoniae raf* operons by various sugars, including raffinose, was investigated, using α-galactosidase activity as a reporter. The *aga* gene exhibits sequence homology to other prokaryotic α-galactosidases,

and this example shows that α-galactosidase activity is encoded by *aga* in *S. pneumoniae*. Hence, the *S. pneumoniae aga* gene, regulated by its promoter P_{AGA}, serves as a naturally-occurring reporter gene for use in the study of induction and regulation of the *S. pneumoniae raf* gene cluster, and can also be used as a reporter gene for the analysis of other potential regulatory sequences in *S. pneumoniae* and other microorganisms.

Cells and Cell Growth

S. pneumoniae strain VSPN3026 was used as wild-type for the experiments described herein. Cells were grown in C+Y Medium (Tomasz (1970) *J. Bacteriol.* **101**:860-871) containing 0.2% (w/v) sucrose and 0.2% (w/v) glucose, and logarithmic phase cells were frozen in 20% (v/v) glycerol, to be used for inoculation. In experiments to identify inducers of galactosidase activity, C+Y medium containing various sugars was used. All sugars were used at a concentration of 0.2% (w/v). Eight ml of medium was inoculated with 200 µl of frozen stock cells, and the culture was grown at 37°C. Growth of the culture was measured by absorbance at 600 nm, using a visible spectrophotometer. When A_{600} reached 0.4, cells were used for experiments.

Measurement of α-galactosidase activity

To prepare cell lysates for measurement of enzyme activity, 1.5 ml samples of cell culture were collected and immediately centrifuged for 5 min. at 14,000 rpm to pellet the cells. The supernatant was discarded and the pellet was resuspended in 0.1 ml of 100 mM sodium phosphate buffer, pH 7.5 containing 0.25% Triton X-100. This mixture was incubated at 37°C for 10 min to lyse the cells.

Alpha-galactosidase activity was measured in a buffer containing 100 mM sodium phosphate, 1 mM $MgCl_2$, 45 mM β-mercaptoethanol, pH 7.5, containing *p*-nitrophenyl-α-D-galactopyranoside (Sigma Chemical, St. Louis, MO) at a final concentration of 0.9 mg/ml. The reaction was initiated by addition of 10 µl cell lysate to 90 µl of reaction buffer and the reaction mixture was incubated at 25°C. Enzymatic activity was monitored by measuring absorbance at 405 nm. A_{405} measurements were taken every 30 sec for 30 min using a Spectramax 250

microtiter plate reader (Molecular Devices, Sunnyvale, CA). Specific activity was calculated using a *p*-nitrophenol standard.

Results are shown in Table 10. As can be seen, low basal levels of α-galactosidase activity were observed in lysates from cells grown with sugars other than raffinose as a sole carbon source. However, a 200-1,000-fold induction of α-galactosidase activity was observed in cells grown on raffinose. Combinations of raffinose and a second sugar gave enzyme levels that were 16-500-fold greater than those obtained with the second sugar alone. Thus, by adjusting the concentration of sugar and/or the combination of sugars in the medium, expression of coding sequences regulated by *raf* regulatory sequences can be modulated over an approximately 1,000-fold range.

Table 10: Galactosidase activity measured in lysates from cells grown on different sugars

Sugar*	Specific Activity (nmol *p*-nitrophenol/min/mg protein)
glucose	12
fructose	20
sucrose	24
galactose	30
lactose	17
maltose	7
raffinose	7,099
fructose + raffinose	4,681
sucrose + raffinose	382
galactose +raffinose	9,714
lactose + raffinose	8,258
maltose +raffinose	3,580

*All sugars are present in the medium at a concentration of 0.2% (w/v)

Construction of mutants in the *aga, rafR and rafS* genes

To characterize the *raf* regulatory system, gene knockouts in the *aga*, *rafR* and *rafS* genes were constructed by the method of Claverys *et al.* (1995) *Gene* **164**:123-128, as follows.

A DNA fragment containing a region of the *aga* gene was amplified by polymerase chain reaction (PCR) using oligonucleotides aga1 and aga2 as primers. *See* Table 11 for the sequences of oligonucleotides. This generated a 320 bp amplification product. The amplified sequence was ligated into pGEM-T Easy (Promega, Madison, WI). The resulting construct was digested with EcoRI, to release an approximately 339 bp fragment containing *aga* sequences. This fragment was inserted into the EcoRI site of pR 326 (Claverys *et al., supra*) to create pR326AGAKO. This plasmid was used to transform VSPN3026 to construct a *S. pneumoniae* strain, VSPN3037, with an insertion of pR326 sequence in the chromosomal *aga* gene, thereby inactivating the *aga* gene.

A DNA fragment containing an internal portion of the *rafR* gene was PCR-amplified using oligonucleotides rafR1 and rafR2 as primers. *See* Table 11 for sequences. This generated a 449 bp amplification product. The amplified sequence was ligated into pCRII (Invitrogen, Carlsbad, CA). The resulting construct was digested with EcoRI, and an approximately 440 bp *rafR*-containing fragment was inserted into the EcoRI site of pR 326 to create pR326RAFRKO. This plasmid was used to transform VSPN3026 to construct a *S. pneumoniae* strain, VSPN3038, with an insertion of pR326 sequence in the chromosomal *rafR* gene, thereby inactivating the *rafR* gene.

A DNA fragment containing an internal portion of the *rafS* gene was PCR-amplified using oligonucleotides rafS1 and rafS2 as primers. *See* Table 11 for sequences. This generated a 454 bp amplification product, which was ligated into pCRII (Invitrogen, Carlsbad, CA). The resulting construct was digested with EcoRI and an approximately 445 bp *rafS*-containing fragment was obtained, which was inserted into the EcoRI site of pR 326 to create pR 326RAFSKO. This plasmid was used to transform VSPN3026 to construct a *S. pneumoniae* strain,

VSPN3039, with an insertion of pR326 sequence in the chromosomal *rafS* gene, thereby inactivating the *rafS* gene.

Table 11: Sequences of Oligonucleotide used for construction of mutations in the *aga, rafR* and *rafS* genes

Oligonucleotide	Sequence (5' → 3')	SEQ ID NO.
aga 1	GCTCAACTTAGTCTGACTTTG	60
aga 2	CAAACACATTCCCAGCATCCTCTG	61
rafR 1	CGCGGATCCTCGAGAAGTTGTCTAGCTCGG	62
rafR 2	CCGGAATTCTAGGAATCACTGGAGGGAAA	63
rafS 1	CCGCGGATCCGCTACAAGTAGTGTGTAGGATGG	64
rafS 2	GCCGGAATTCAATCCTACCAAGCTGTCTACC	65

Characterization of mutants

The *aga, rafR and rafS* mutant strains were tested for growth and for α-galactosidase activity, when provided with raffinose, sucrose or a mixture of raffinose and sucrose as carbon source. Mutant strains were grown in C+Y medium containing different carbon sources (as indicated in Table 12) and growth was monitored by absorbance of cultures at 600 nm, measured by spectrophotometry. When cultures reached an A_{600} of 0.4, cells were collected and assayed for α-galactosidase activity as described *supra*. Assay results are shown in Table 12.

Strains with a mutation in *aga* were unable to grow on raffinose. The inability of an *aga* mutant strain to grow on raffinose confirmed that mutation had occurred in a gene necessary for raffinose metabolism and, since raffinose is an α-galactoside, is consistent with the inactivation of an α-galactosidase.

The *aga* mutant strains grew when either sucrose or sucrose + raffinose were provided as carbon sources, but exhibited non-measurable levels of α-galactosidase activity under both of these conditions. *See* Table 12. These results indicate that the α-galactosidase activity observed in wild-type cells grown on raffinose is provided by the product of the *aga* gene. Taken together, the

results indicate that expression of the *aga* gene is activated by raffinose; *i.e.*, raffinose is an inducer of *aga*.

A strain harboring a mutation in the *rafR* gene is able to grow on raffinose, but induced levels of α-galactosidase in this strain are seven-fold lower than in wild-type cells. *See* Table 12. Thus, *rafR* function is required for maximal induction of *aga* activity. These results are those expected if the *rafR* gene product acts as an activator of the *aga* gene, and are consistent with the presence, in the amino acid sequence of the *rafR* gene product, of the AraC family signature sequence. They are also consistent with the high degree of homology between the RafR amino acid sequence and the sequences of other transcriptional activator proteins.

Strains with *rafR* mutations grow slowly when raffinose is provided as a sole carbon source, as expected if RafR is an inducer of the *raf* metabolic operon. However, *rafR* mutants also grow more slowly than wild-type when sugars other than raffinose are provided as a carbon source. This suggests that RafR has additional regulatory targets outside the *raf* metabolic operon, and that the RafR protein may have additional regulatory functions beyond those that are related to raffinose metabolism. Products of non-*raf* genes that are regulated by RafR may serve as potential targets for drug discovery.

Strains carrying mutations in *rafS* that were grown in the presence of raffinose (*i.e.*, either raffinose alone or raffinose + sucrose) express higher levels of α-galactosidase activity than wild-type cells. These results are consistent with the *rafS* gene product being a negative regulator of *aga* expression.

An alternative interpretation of these results is possible if, for example, inactivation of *rafR* has a polar effect on *rafS*, such that the activity detected in *rafR* mutants reflects absence of both RafR and RafS function. If this were the case, the *rafR* gene product could be an activator, or the combined activity of the *rafR* and *rafS* gene products could provide activator function.

Table 12: α-galactosidase activity in mutant strains grown on different carbon sources

Strain	Specific Activity (nmol *p*-nitrophenol/min/mg protein)		
	Sucrose	Raffinose	Sucrose + Raffinose
VSPN3026 (wild-type)	4.8	2,538	137
VSPN3037(*aga*⁻)	0.1	ND (no growth)	0.1
VSPN3038(*rafR*⁻)	3.9	346	32
VSPN3039 (*rafS*⁻)	8.3	3,756	403

While the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be apparent to those skilled in the art that various changes and modifications may be practiced without departing from the spirit of the invention. Therefore the foregoing descriptions and examples should not be construed as limiting the scope of the invention.

CLAIMS

What is claimed is:

1. A cell which expresses a gene involved in an essential cellular process, wherein expression of the gene can be regulated over a range of levels, and further wherein the range includes a low basal expression level.

2. A cell according to claim 1, wherein the basal expression level is less than about 50% of wild type.

3. A cell according to claim 1, wherein the basal expression level is insufficient to support cell growth.

4. A cell according to claim 1, wherein the gene encodes a polypeptide.

5. A cell according to claim 1, wherein the gene encodes an untranslated RNA molecule.

6. A cell according to claim 4, wherein the polypeptide is selected from the group consisting of a structural protein, an enzyme, a receptor, an intracellular signaling molecule and a cellular adhesion molecule.

7. A cell according to claim 1, wherein the cellular process is selected from the group consisting of replication, recombination, DNA repair, transcription, translation, protein processing, protein export, cell wall biosynthesis, cell membrane synthesis, lipid metabolism, protein metabolism, energy metabolism, cell division, drug resistance and virulence.

8. A cell according to claim 1, wherein expression of the gene is regulated by fusion of the gene to a heterologous regulatory element.

9. A cell according to claim 4, wherein expression of the polypeptide is regulated by a heterologous regulatory element that has been fused to a sequence which encodes the polypeptide or a fragment thereof.

10. A cell according to claim 9, wherein the regulatory element is the P_{BAD} promoter.

11. A cell according to claim 10, wherein expression is regulated by adjusting the concentration of L-arabinose or L-ribose in the cellular growth medium.

12. A cell according to claim 11, wherein further regulation is achieved by adjusting the concentration of glucose or other carbon source in the cellular growth medium.

13. A cell according to claim 9, wherein the regulatory element is selected from the group consisting of the *lac* promoter, the *trp* promoter, the *tac* promoter, the *gal* promoter, the *lpp* promoter, the *phoA* promoter, the T7 promoter, the T3 promoter, the SP6 promoter, the λp_R promoter, the λp_L promoter, and the *tet* promoter.

14. A cell according to claim 9, wherein the heterologous regulatory element is regulated by a member of the AraC/XylS family.

15. A cell according to claim 9, wherein the heterologous regulatory element is regulated by a two-component regulatory system.

16. A cell according to claim 1, wherein the cell is a microorganism.

17. A cell according to claim 16, wherein the cell is a prokaryotic cell.

18. A cell according to claim 17, wherein the cell is Gram-positive.

19. A cell according to claim 17, wherein the cell is Gram-negative.

20. A cell according to claim 1, wherein the cell is a eukaryotic cell.

21. A cell according to claim 20, wherein the cell is a yeast cell.

22. A cell according to claim 20, wherein the cell is a fungal cell.

23. A cell according to claim 20, wherein the cell is a plant cell.

24. A cell according to claim 20, wherein the cell is a mammalian cell.

25. A cell according to claim 24, wherein the cell is a human cell.

26 A cell according to claim 25, wherein the cell is malignant.

27. A cell according to claim 26, wherein the cell is resistant to a therapeutic.

28. A cell according to claim 17, wherein the cell is resistant to a therapeutic.

29. A cell according to claim 20, wherein the cell is resistant to a therapeutic.

30. A method for identifying a compound that affects an essential cellular process, the method comprising the steps of:

(a) providing a cell according to claim 1,

(b) exposing the cell to the compound; and

(c) assaying cell viability.

31. The method according to claim 30, wherein the cell is cultured in a liquid medium and the cell is exposed to the compound by addition of the compound to the culture medium.

32. The method according to claim 30, wherein the cell is cultured on a solid medium and the cell is exposed to the compound by application of the compound to the solid medium.

33. The method according to claim 30, wherein cell viability is assayed by measuring cell growth.

34. The method according to claim 33, wherein cell growth is determined by a measurement selected from the group consisting of vital staining, cell counting, light scattering, incorporation of macromolecular precursor, fluorescence-activated cell sorting and reporter gene expression.

35. A method for determining the target of a compound, the method comprising the steps of:

(a) providing a library of cells according to claim 1, wherein, in each member of the library, the expression of a different gene product is regulated, and wherein, in the cells comprising the library, a variety of different gene products are regulated;

(b) exposing the library to the compound; and

(c) assaying cell growth;

wherein, if the growth of a member of the library is negatively affected, the gene product regulated in that member is the target.

36. A method for identifying a gene that is involved in an essential cellular process, the method comprising the steps of:

(a) constructing a fusion between a heterologous regulatory element and a coding sequence, wherein the heterologous regulatory element allows expression of the coding sequence over a range of levels, and further wherein the range includes a low basal expression level,

(b) exposing a cell containing the fusion to a concentration of a test compound, and

(c) assaying cell viability in the presence of the test compound.

37. The method according to claim 36, wherein the fusion is constructed *in vitro* and is introduced into a cell.

38. The method according to claim 36, wherein the fusion is constructed *in vivo* within the cell.

39. The method according to claim 36, wherein cell viability is assayed by measuring cell growth.

40. The method according to claim 36, wherein viability is assayed at more than one level of expression of the coding sequence.

41. The method according to claim 36, wherein the coding sequence encodes an untranslated RNA molecule

42. The method according to claim 36, wherein the coding sequence encodes a polypeptide or a fragment thereof

43. The method according to claim 36, wherein viability is assayed at more than one concentration of the test compound.

44. The method according to claim 40, wherein viability is assayed at more than one concentration of the test compound.

45. The method according to claim 35, wherein the regulatory element is the P_{BAD} promoter.

46. The method according to claim 45, wherein expression is regulated by adjusting the concentration of L-arabinose or L- ribose in the cellular growth medium.

47. The method according to claim 46, wherein further regulation is achieved by adjusting the concentration of glucose or other carbon source in the cellular growth medium.

48. The method according to claim 36, wherein the regulatory element is selected from the group consisting of the *lac* promoter, the *trp* promoter, the *gal* promoter, the T7 promoter, the T3 promoter, the SP6 promoter, the λp_R promoter, the λp_L promoter and the *tet* promoter.

49. The method according to claim 36, wherein the heterologous regulatory element is regulated by a member of the AraC/XylS family.

50. The method according to claim 36, wherein the heterologous regulatory element is regulated by a two-component regulatory system.

51. The method according to claim 36, wherein the basal expression level is less than about 50% of wild type.

52. The method according to claim 42, wherein the coding sequence encodes a polypeptide selected from the group consisting of a structural protein, an enzyme, a receptor, an intracellular signaling molecule and a cellular adhesion molecule.

53. The method according to claim 36, wherein the cellular process is selected from the group consisting of replication, recombination, DNA repair, transcription, translation, protein processing, protein export, cell wall biosynthesis, cell membrane synthesis, lipid metabolism, protein metabolism, energy metabolism, cell division, drug resistance and virulence.

54. The method according to claim 36, wherein the cell is cultured in a liquid medium and the cell is exposed to the compound by addition of the compound to the culture medium.

55. The method according to claim 36, wherein the cell is cultured on a solid medium and the cell is exposed to the compound by application of the compound to the solid medium.

56. The method according to claim 36, wherein cell growth is determined by a measurement selected from the group consisting of vital staining, cell counting, light scattering, incorporation of macromolecular precursor, fluorescence-activated cell sorting and reporter gene expression.

57. A method for identifying a gene responsible for conferring sensitivity to a test compound, the method comprising identifying a gene according to the method of claim 36.

58. A method for identifying a gene responsible for conferring resistance to an antibiotic, the method comprising identifying a gene according to

claim 36 wherein the test compound is an antibiotic and the essential cellular function is antibiotic resistance.

59. A method for identifying a gene responsible for virulence, using the cell of claim 1, wherein the essential cellular process is involved in virulence.

60. A method for identifying a polypeptide that is involved in an essential cellular process, the method comprising identifying a gene according to claim 36 and determining the polypeptide from the sequence of the gene.

61. A method for identifying an RNA that is involved in an essential cellular process, the method comprising identifying a gene according to claim 36 and determining the identity of the RNA from the sequence of the gene.

62. A cell according to claim 9, wherein the heterologous regulatory element is selected from the group consisting of the *mal Px* promoter and the *mal Pm* promoter.

63. A cell according to claim 62, wherein expression is regulated by adjusting the concentration of maltose in the growth medium.

64. A cell according to claim 63, wherein the cell is grown in a minimal medium.

65. The method according to claim 36, wherein the heterologous regulatory element is selected from the group consisting of the *mal Px* promoter and the *mal Pm* promoter.

66. The method according to claim 65, wherein expression is regulated by adjusting the concentration of maltose in the growth medium.

67. The method according to claim 66, wherein the cell is grown in a minimal medium.

68. The modified *mal Px* promoter according to Figure 1, wherein the sequence GGA at approximately -80 is converted to GCG.

69. A cell according to claim 9, wherein the heterologous regulatory element is selected from the group consisting of *raf* P_{AGA}, *raf* P_R *and raf* P_E.

70. A cell according to claim 69, wherein expression is regulated by adjusting the concentration of raffinose in the growth medium.

71. A cell according to claim 69, wherein expression is regulated by adjusting the concentration of sucrose in the growth medium.

72. The method according to claim 36, wherein the heterologous regulatory element is selected from the group consisting of *raf* P_{AGA}, *raf* P_R *and* *raf* P_E.

73. The method according to claim 72 wherein the heterologous regulatory element is regulated by adjusting the concentration of raffinose in the growth medium.

74. The method according to claim 72 wherein the heterologous regulatory element is regulated by adjusting the concentration of sucrose in the growth medium.

75. The method according to claim 30, wherein the cell is hypersusceptible to the compound.

76. The method according to claim 75, wherein hypersusceptibility is due to a mutation in a cellular gene.

77. The method according to claim 76, wherein the cellular gene encodes a component of an efflux pump.

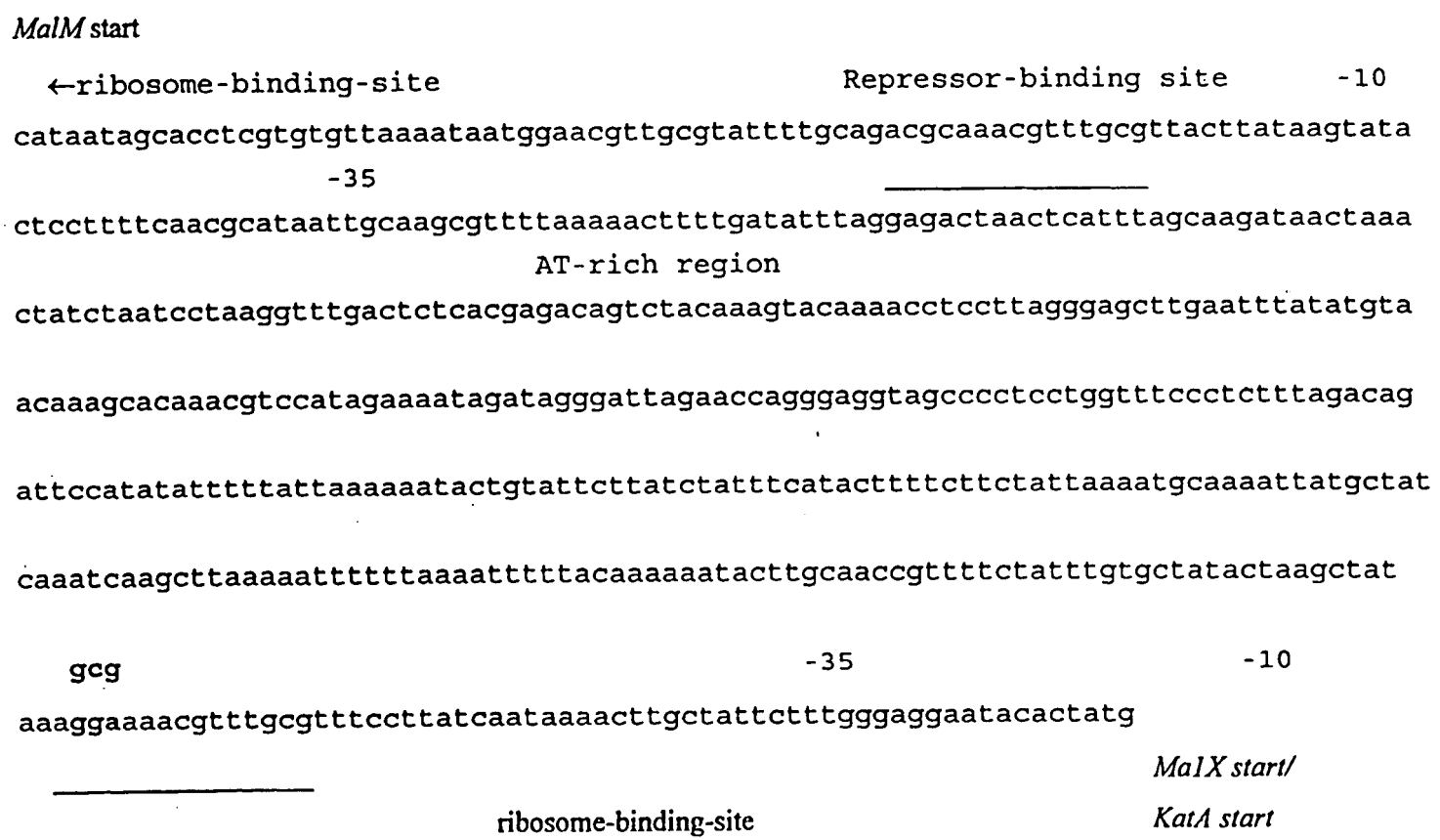
MalM start
←ribosome-binding-site
Repressor-binding site
-10
cataatagcacctcgtgtgttaaaataatggaacgttgcgtattttgcagacgcaaacgtttgcgttacttataagtata
-35
ctccttttcaacgcataattgcaagcgttttaaaaacttttgatatttaggagactaactcatttagcaagataactaaa
AT-rich region
ctatctaatcctaaggtttgactctcacgagacagtctacaaagtacaaaacctccttagggagcttgaatttatatgta
acaaagcacaaacgtccatagaaaatagatagggattagaaccagggaggtagcccctcctggtttccctctttagacag
attccatatattttattaaaaaatactgtattcttatctatttcatacttttcttctattaaaatgcaaaattatgctat
caaatcaagcttaaaaatttttttaaaatttttacaaaaaatacttgcaaccgttttctatttgtgctatactaagctat
gcg
-35
-10
aaaggaaaacgtttgcgtttccttatcaataaaacttgctattctttgggaggaatacactatg
MalX start/
KatA start
ribosome-binding-site
Repressor-binding site

FIG._1

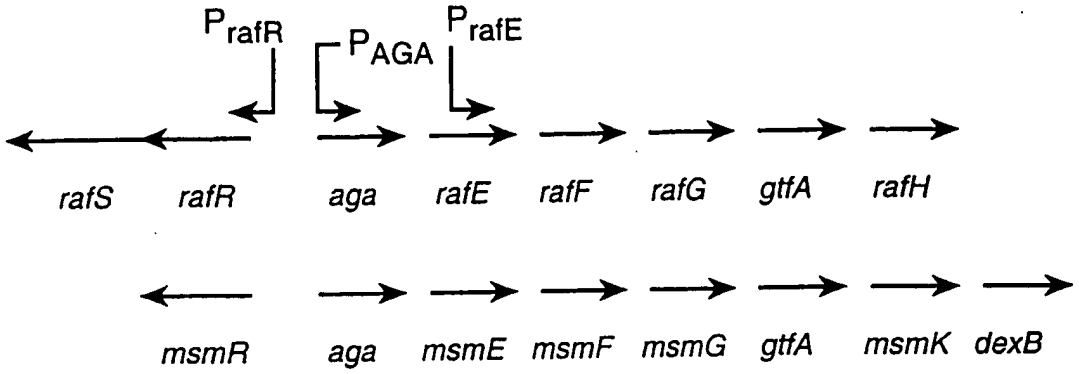
P rafR
P AGA
P rafE
rafS
rafR
aga
rafE
rafF
rafG
gtfA
rafH
msmR
aga
msmE
msmF
msmG
gtfA
msmK
dexB

FIG._2

```
ATTTTACTTC CAACTATTGA GAGAAATTTC GCCACTATTT AGCCAGATTT CTTTTCCGTT 60
ATCACATTGA ACTAAAAGTT TTCCATTTTC TGAGATGTCT TTAGCAAGTC CCTTGTAGTC 120
TTTTTGCTCT AGTGTGAAAG TGACTTCTTT TCCTAGAATG AATGACTGTT TTTTGTATAG 180
GTATAATAGC TCTTCTGCTG GTGTTTCGAA GAAAGCACGC CAGATTTCTA TGATCAATTC 240
ATTCCTTGTT ATAGGAGCTG TAGCTTTAAA TAAGCTGGCA GCTTTTTCTT TTAATTCCTG 300
AGGGAAGTCT TTAATAGTGA AATTGATCCT ACTCCAATAA TGATATCTGT GACTAAGCCA 360
GTTTCTACAG AGGTCATTGC TTCAGTAAGG ATTCCTCCAA TTTTATGATT GTTTAGATAG 420
ATATCATTGA CCCATTTTAT ATCGACATCT ATTAAAGTTA GGTTCTTAAT GGCTTTGTAG 480
ACAGCTCCAG CTACAAGTAG TGTGTAGGAT GGTAATTTGT CATAGGGGAG ATTTGGTTTA 540
AGATGGAGTG TCATATAAAT TCCACCTTGT GGTGAGTAGA AGGAACGTTG AAAACGGCCT 600
CGGCCTGCTG TTTGATAGGA AGCTAGATAG AGAGTATTTG CTTCATGGCC TAAATCAATT 660
GCTTCTTTTG CATCTAGTTG TGTTGATTTT GTTTCGGGTT TAAAGCTGAC TTTAATTGGA 720
AGATTTTCTT CTAGAATCTC TGGAAGAATA AGGTCACCAT TCATTAGTTT ATATCCTCTA 780
TTTTTGATAC TATTAATTTC AATGCCTTCT TGTTCTAGTC GCTTGATAGC TTTCCAAATT 840
GATGTTCGGC TTAGGGATAG TTTTTCTGCG ATTTTTTCTC CGCTGATATA GTCGGTTTCT 900
TTAGATAGGA TTTGGTAGAC AGCTTGGTAG GATTTCATAA TGTTGCCTTT CTCACTAGTT 960
GGTATTGAGA GTATTCTTTT CTTGTATGAC TTGGAGACTG ATTAAAGTAT TGTTTATAAG 1020
CTTTCGAAAA ATGGAGTGGA TCTGAAAAAC CTACCGAGTA TGCAATTACC TTGATGGACT 1080
CTTGGGTATT TTCGAGAAGT TGTCTAGCTC GGTGCATTCG AACGTAGAGT AGGTATTCTT 1140
TGGGTGATAA GGTATTAAAT TCTTTGAATA CGCTTGATAA GTAGCTTCTG TGAACGGATA 1200
GTTCTTTTGC TAAATCTTGA ATTGTAAGTG ATTGAGGATA GTGGCTATCA ATTAATCGTT 1260
TGCATTCAAG ATAGAGTTGG TGGGTTGATG AAATATTCTT TTTTTTCTGA TTGGGAGCAA 1320
TAGTTCCCAG ATGAAACATC AGTTCATGAA GTTGTCCCAT GATATGGAGT TGAGCTAATT 1380
CACTTGATTT TGTAATCTGA GCGAAGCGGA CAATGTCTGA GATGAGTTTT GCAGTAGTCT 1440
GGGTATGACA AGTTTCAGAT TGGATGAGAT AGGATTGATC AGAAATTTGA GAAAGAGCAA 1500
AATAATCAGG GGATTTCCCT CCAGTGATTC CTAACCAGTA GTAGGCCCAA GGTTCTTTAC 1560
TATCTGCTTG ATAAAAGGTT AGTTCCTCTG GTTTTAATAG AAAGAAATCT CCTTCTTTTA 1620
AATCAACAAT TTTACCCTTG TAATGAAATT TTCCTTGTCC TTTACTAATG TAATGTAGGA 1680
CGTATGTATC ACGAATGGCT GGACCAAAAG AGTAATTAGG TGTGCATTCC TCATATCCAT 1740
AAAAGCTTAG GGCAAGGTCA ATTGTTCCAG TCTGGTATTC TGAAAAAACT AGCATGTGCT 1800
ACCTCCTACC TAACATTTTA CCATATTCTT GAGACATTTT TCTATTTTGG AAAGCGATTT 1860
CAGGTGATAA AATATAATCA ATAAAGTGAT GAGGTGAAGT AAATGGGAGT TAGGATAGAG 1920
AATAATCTAT TTTATGTTGA GAGTAAAAAT CTAAGTTTGA TTATTGAAAA TCGAAATGGC 1980
TACTTACTTT TGAAACATTT AGGAAAGACT ATTAAGAACT ATAAAGGTTC CAATAGTGTT 2040
TATGAACGAG ACCATGCCTT TTCAGGAAAT CCAACGGCTA CTAATCGAAC CTTTAGTTTA 2100
GATACTCAGC GACAGATTTT TGGACAACAT GGCTTAGGAG ATTTTAGGAA ACCAACTATA 2160
CAGGTTCAGC ATAGTGTAAC TGAAGTAACA GACTTTCGAT TTGTAGAAGC AAAGATTTTA 2220
AAAGGTCAGA ATGGTCCACA GGGCTTACCT TCTCCACATA GCATGGACGA TACAGAGACT 2280
CTTGTCTTAA TGTTAGAAGA TTCTAAGGCT CAACTTAGTC TGACTTTGTA TTATACTACT 2340
TTTAATAATG ATGCGACTAT TGCTAGCTAC AGTAAATTAG ATAATAATAG TAATCAGGAA 2400
GTTGTCATCC ATAAAGATTT TTCTTTTATG GCTGATTTTC CAGCTGCAGA TTACGAAATA 2460
GTAACTCTGC AGGGTGCTTA TGCTCGTGAA AAGACTGTTA GACGTCAACA GGTAGAACAA 2520
GGAATCTTTT CGATTAGTTC AAACCGAGGT GCTTCTGGTC ATGCTCAAAC ACCAGCTCTT 2580
CTACTATGCG AACAAGGAGT CACAGAGGAT GCTGGGAATG TGTTTGCTAT TCAACTAATG 2640
TATAGTGGCA ACTTTGAAGC TTTTGTTCAA AAAAATCAAT TGAATGAAGT TCGGGTGGCT 2700
ATTGGCATTA ATCCAGAAAA CTTTTCTTGG AAGTTAGCCC CTGAGGAATA CTTTGAAACA 2760
CCGGTAGCTT TAGTGACTCA TTCAGATCAG GGATTAACTG GTATTAGTCA TGAAAGTCAG 2820
AATTTTGTAC TGAAGCACAT TATTCTAAGT GAATTTTCTA AAAAAGAACG TCCAATTCTA 2880
ATCAATAACT GGGAAGCTAC TTACTTTGAC TTTCAGAGAG AAAAACTGTT AGAGTTAGCA 2940
```

FIG._3-1

```
GATGAAGCTA AGAAAGTTGG CATTGAACTT TTTGTATTAG ATGATGGTTG GTTTGGCAAT 3000
CGTTTTGATG ATAATCGTGC TTTAGGTGAT TGGGTTGTTA ATGAGGAAAA ACTGGGTGGA 3060
AGTCTAGAAA GTCTGATTTC AGCTATCCAT GAAAGAGGTT TGCAGTTTGG ACTTTGGTTA 3120
GAACCCGAAA TGATTTCTGT AGATAGTGAT TTGTATCGTC AACATCCTGA CTGGGCTATT 3180
CAGGTTCCTG ATTATGAGCA TACTTATTCT CGGAATCAAT TAGTACTTAA TCTTGCCAAT 3240
CCTCAGGTAG TAGAATACTT GAAAAGTGTC TTAGATCAAC TCCTATCTTA TCATGAGATT 3300
GATTACATTA AATGGGATAT GAACCGCAAT ATCACTAAGC TAGGGAATGG ATTAACTTAT 3360
CTAGAGACAC AGATGCAATC TCATCAGTAC ATGCTGGGAC TTTACGAACT CGTTTCTTAT 3420
CTGACAGAGA AGCACAGCCA TATTCTCTTT GAGTCCTGCT CTGGTGGTGG TGGACGAAAT 3480
GATCTTGGTA TGATACGCTA TTTCCCACAA GTCTGGGCTA GTGATAATAC TGATGCCATT 3540
GCACGTTTAC CAATTCAATA CGGTTCATCC TATCTCTATC CAACCATTTC TATGGGGGCT 3600
CATGTGTCAG CAGTACCGAA TCATCAGATG GGACGAATGA CACCATTAGA AACACGTGGC 3660
CATGTAGCAA TGATGGGAAA TTTGGGCTAT GAGCTTGATT TGACAAATTT ATCAGATGAA 3720
GAGAAAGCTA CGATTGCTAA TCAGGTGAAC TTGTATAAAG AATTACGACC AGTAGTTCAG 3780
TTAGGACAGC AGTATAGACT AATTAATCCT GATACTGCAT CCAATGAAGC AGCTGTACAA 3840
TTTAATTACG GAAATCAAAC GATTGTAACC TACGTTCGCG TTTTATCTGT TGTAGAGACC 3900
ATGGAAACAA CTTTAAAATT AAAAGATTTG GATGAAGAGG GACTATATAA ATTACAGGAA 3960
AATGGCGAAG TTTACTCAGG TGCAGAACTA ATGTATGCGG GCTTAACTGT TATCTTATCC 4020
CAAGGAGATT TTTTGAGTAG ACAGTATATT TTTAGAAAAC TATGA                  4065
```

FIG._3-2

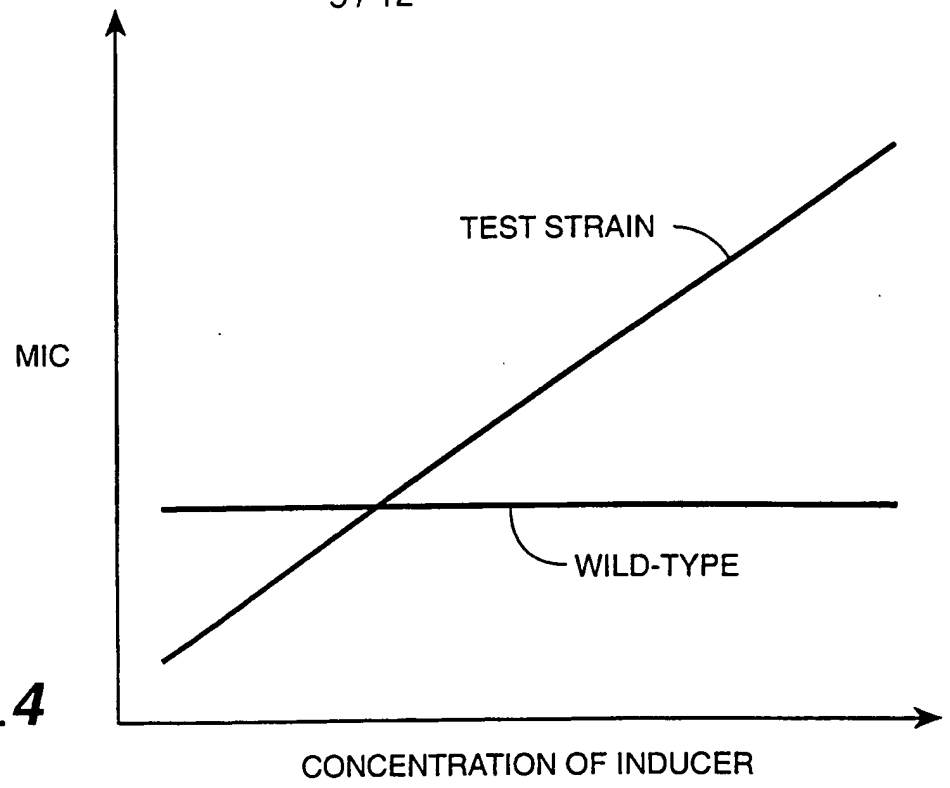
TEST STRAIN
MIC
WILD-TYPE
CONCENTRATION OF INDUCER

FIG._4

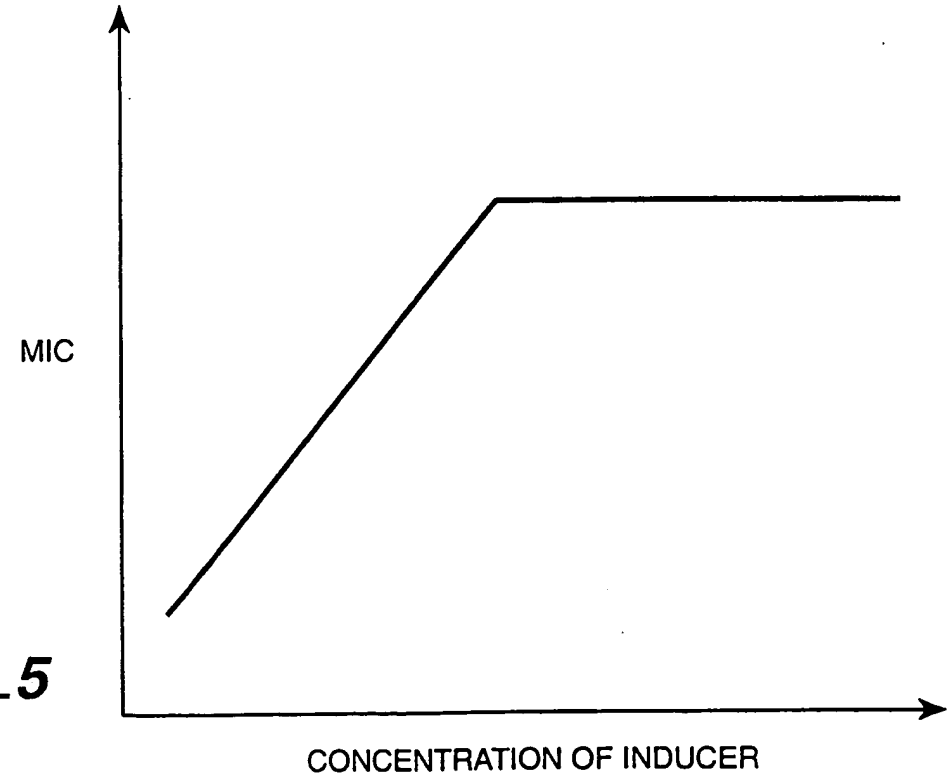
MIC
CONCENTRATION OF INDUCER

FIG._5

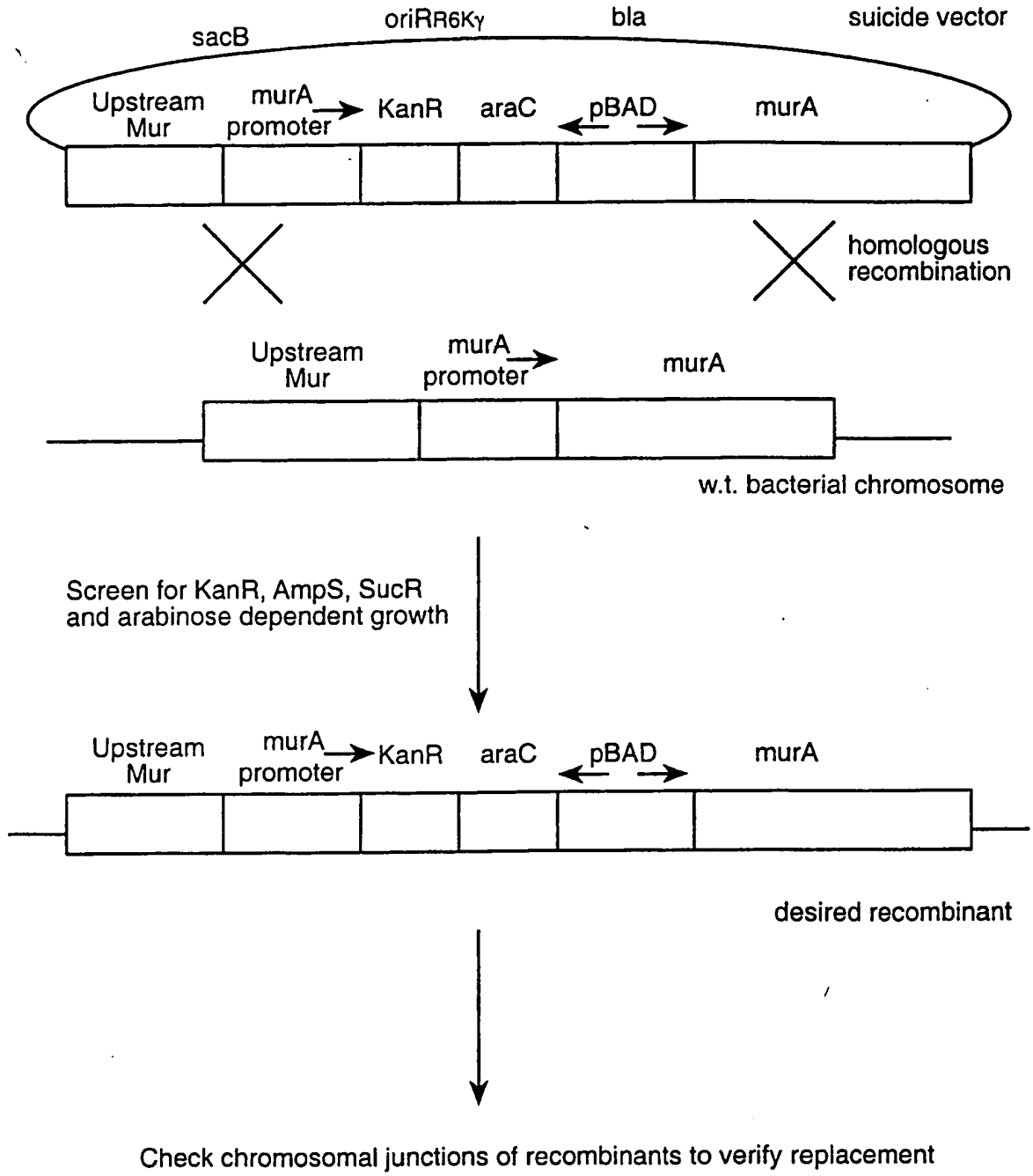
sacB
oriRR6Kγ
bla
suicide vector
Upstream Mur
murA promoter
KanR
araC
pBAD
murA
homologous recombination
Upstream Mur
murA promoter
murA
w.t. bacterial chromosome
Screen for KanR, AmpS, SucR and arabinose dependent growth
Upstream Mur
murA promoter
KanR
araC
pBAD
murA
desired recombinant
Check chromosomal junctions of recombinants to verify replacement

FIG._6

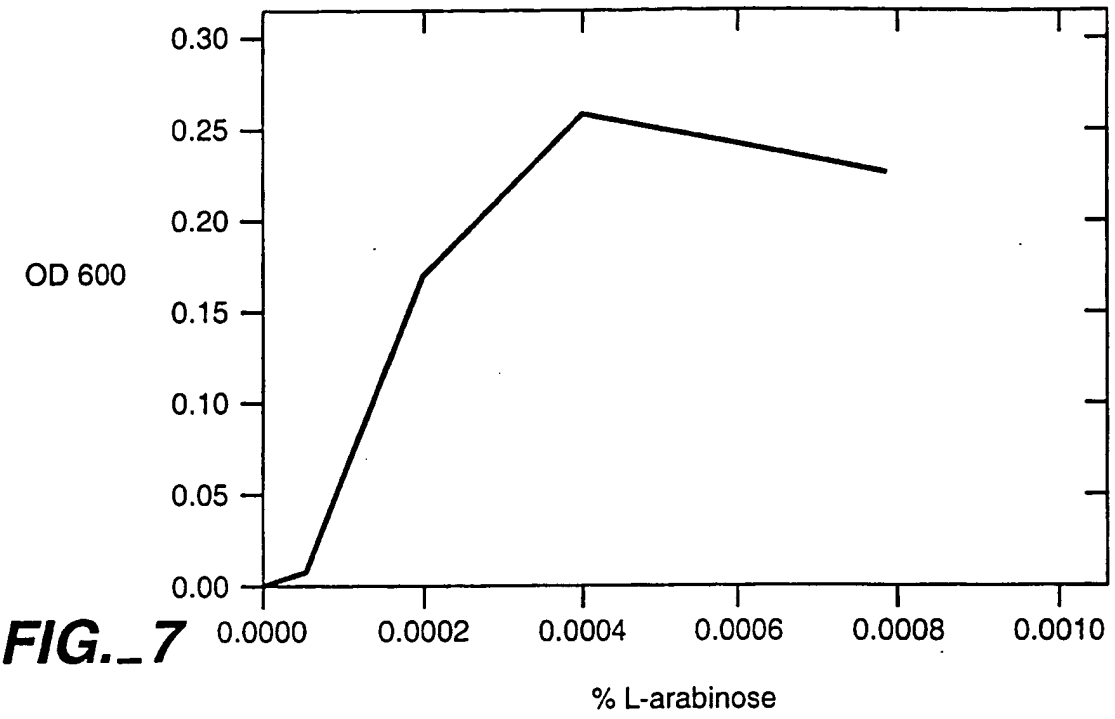
0.30
0.25
0.20
0.15
0.10
0.05
0.00
OD 600
0.0000
0.0002
0.0004
0.0006
0.0008
0.0010
% L-arabinose

FIG._7

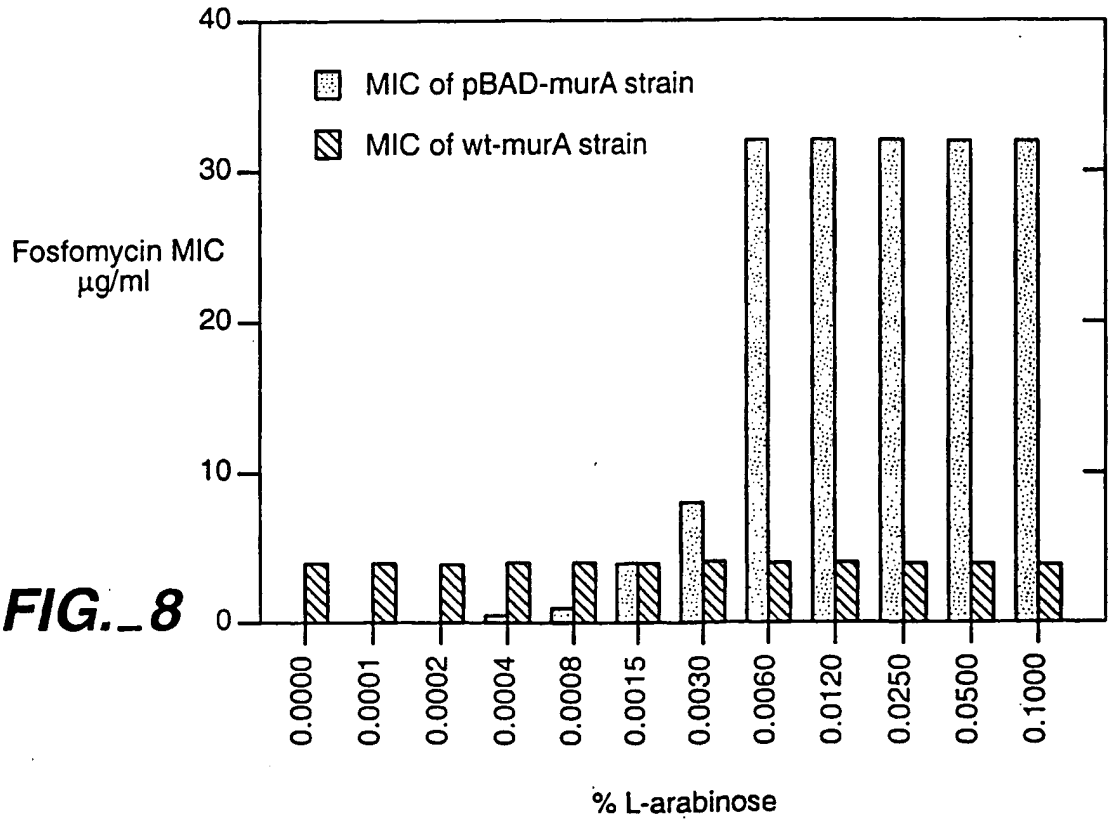
40
30
20
10
0
MIC of pBAD-murA strain
MIC of wt-murA strain
Fosfomycin MIC
µg/ml
0.0000
0.0001
0.0002
0.0004
0.0008
0.0015
0.0030
0.0060
0.0120
0.0250
0.0500
0.1000
% L-arabinose

FIG._8

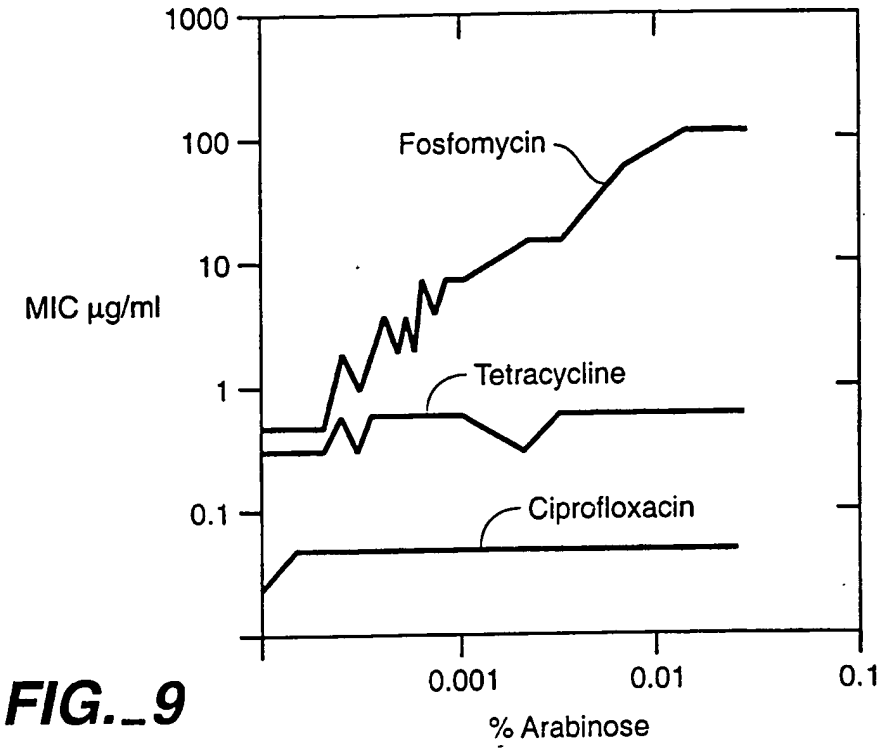
1000
100
10
1
0.1
MIC µg/ml
Fosfomycin
Tetracycline
Ciprofloxacin
0.001
0.01
0.1
% Arabinose

FIG._9

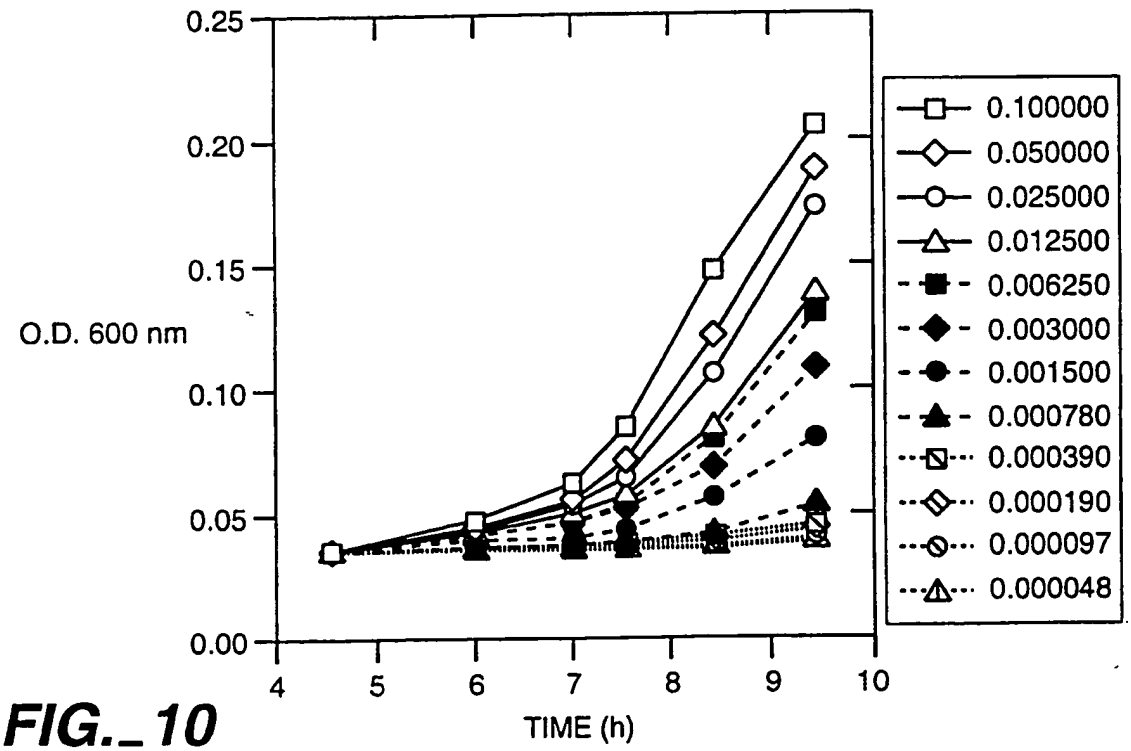
0.25
0.20
0.15
0.10
0.05
0.00
O.D. 600 nm
4
5
6
7
8
9
10
TIME (h)
0.100000
0.050000
0.025000
0.012500
0.006250
0.003000
0.001500
0.000780
0.000390
0.000190
0.000097
0.000048

FIG._10

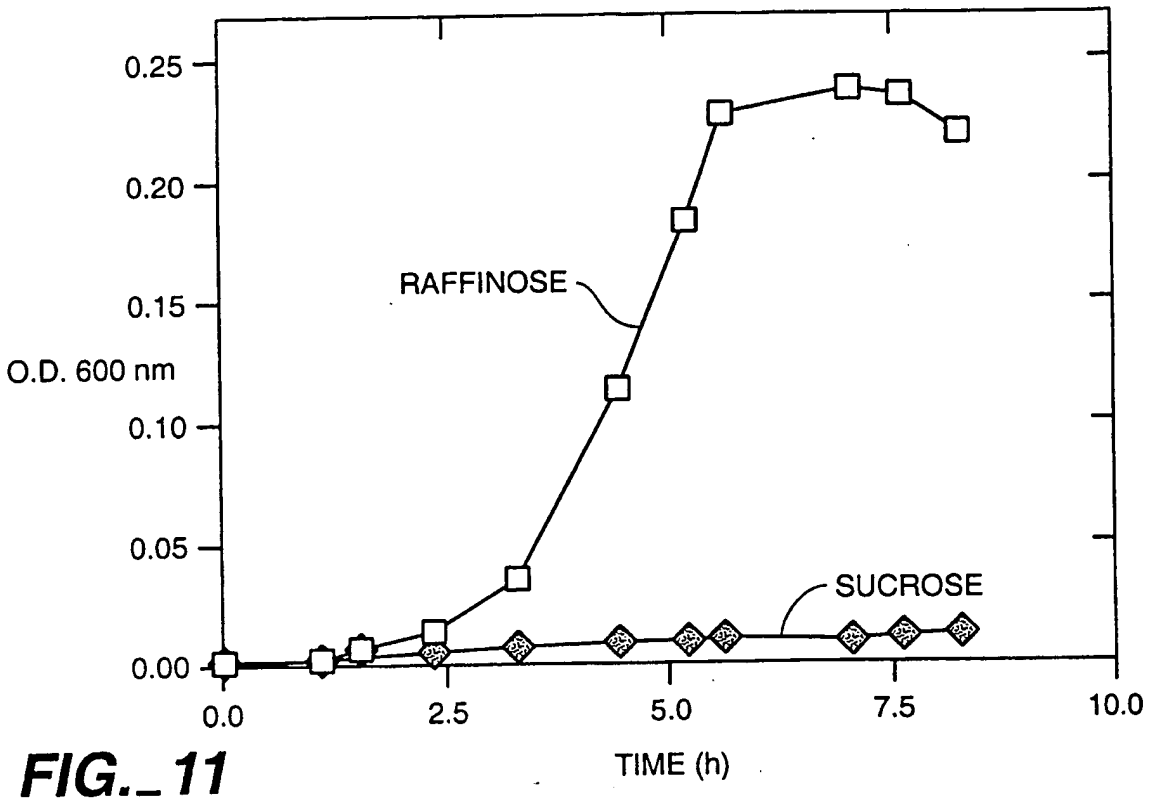
0.25
0.20
0.15
0.10
0.05
0.00
O.D. 600 nm
RAFFINOSE
SUCROSE
0.0
2.5
5.0
7.5
10.0
TIME (h)

FIG._11

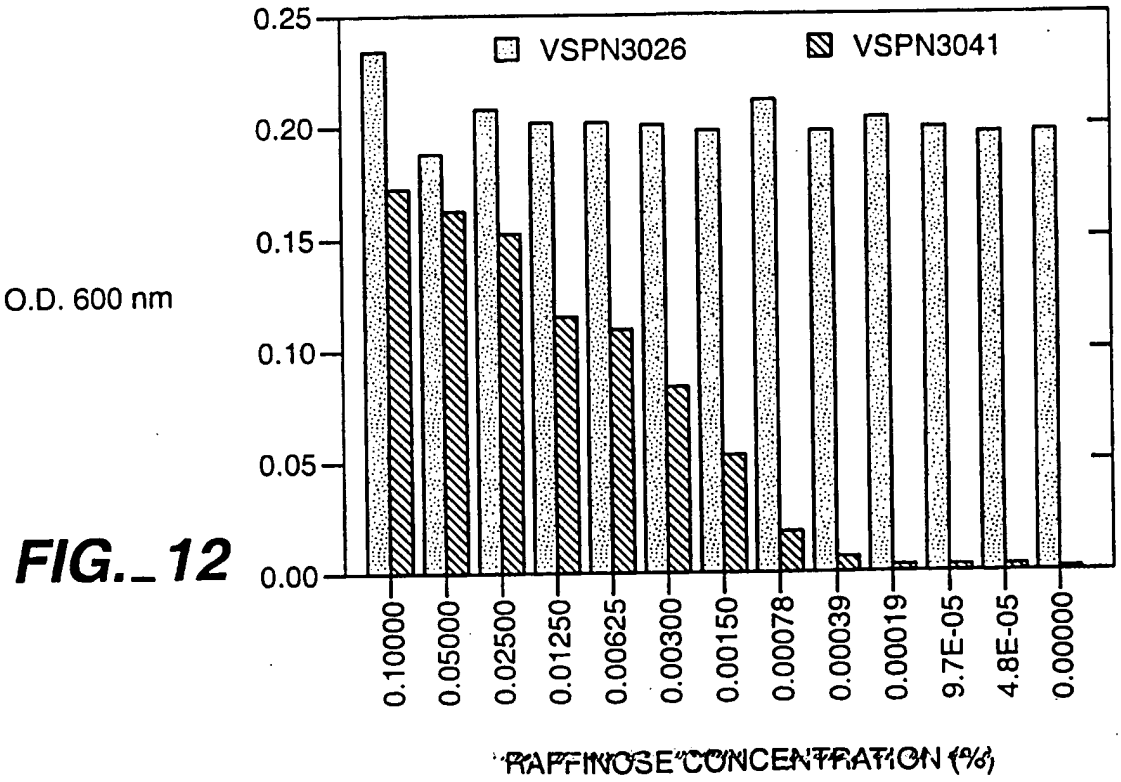
0.25
0.20
0.15
0.10
0.05
0.00
O.D. 600 nm
VSPN3026
VSPN3041
0.10000
0.05000
0.02500
0.01250
0.00625
0.00300
0.00150
0.00078
0.00039
0.00019
9.7E-05
4.8E-05
0.00000
RAFFINOSE CONCENTRATION (%)

FIG._12

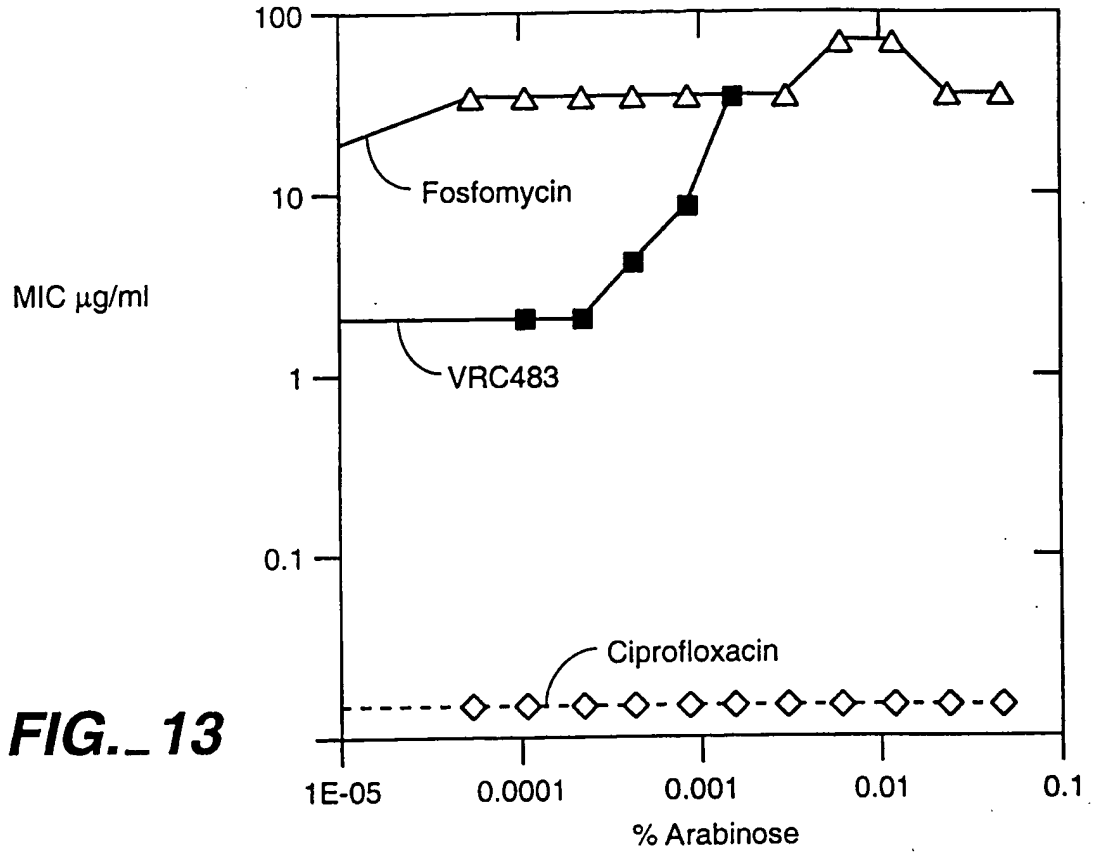
100
10
1
0.1
MIC µg/ml
Fosfomycin
VRC483
Ciprofloxacin
1E-05
0.0001
0.001
0.01
0.1
% Arabinose

FIG._13

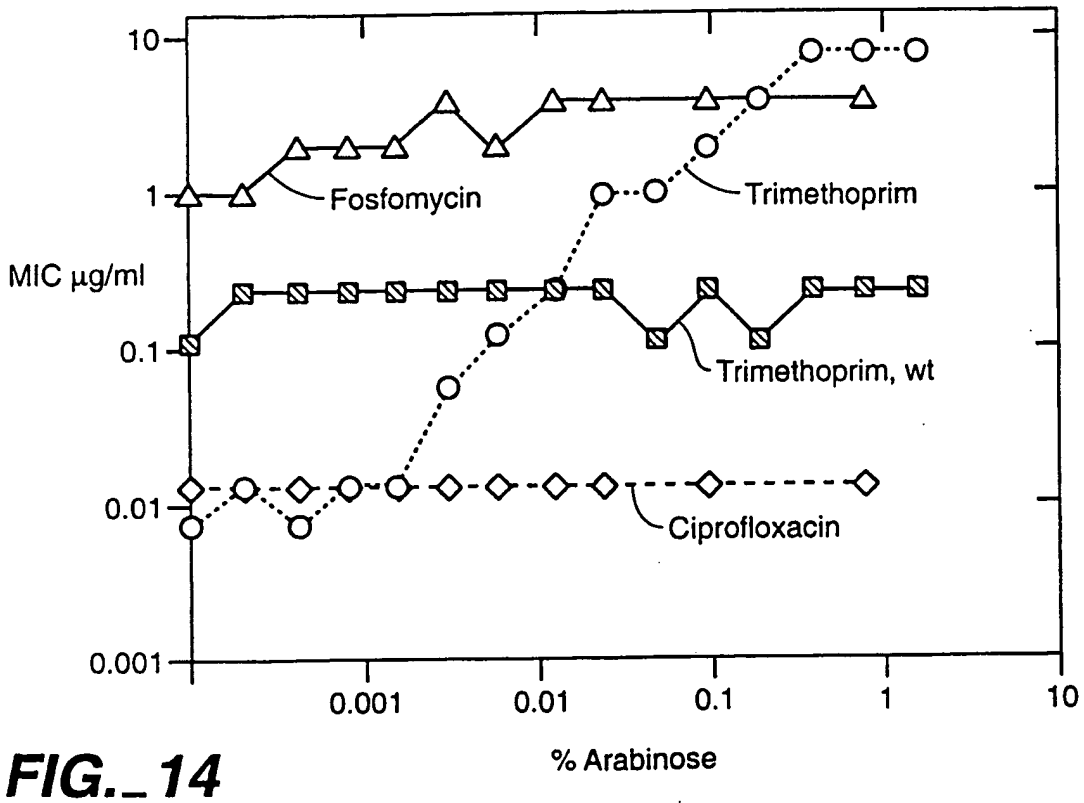
10
1
0.1
0.01
0.001
MIC µg/ml
Fosfomycin
Trimethoprim
Trimethoprim, wt
Ciprofloxacin
0.001
0.01
0.1
1
10
% Arabinose

FIG._14

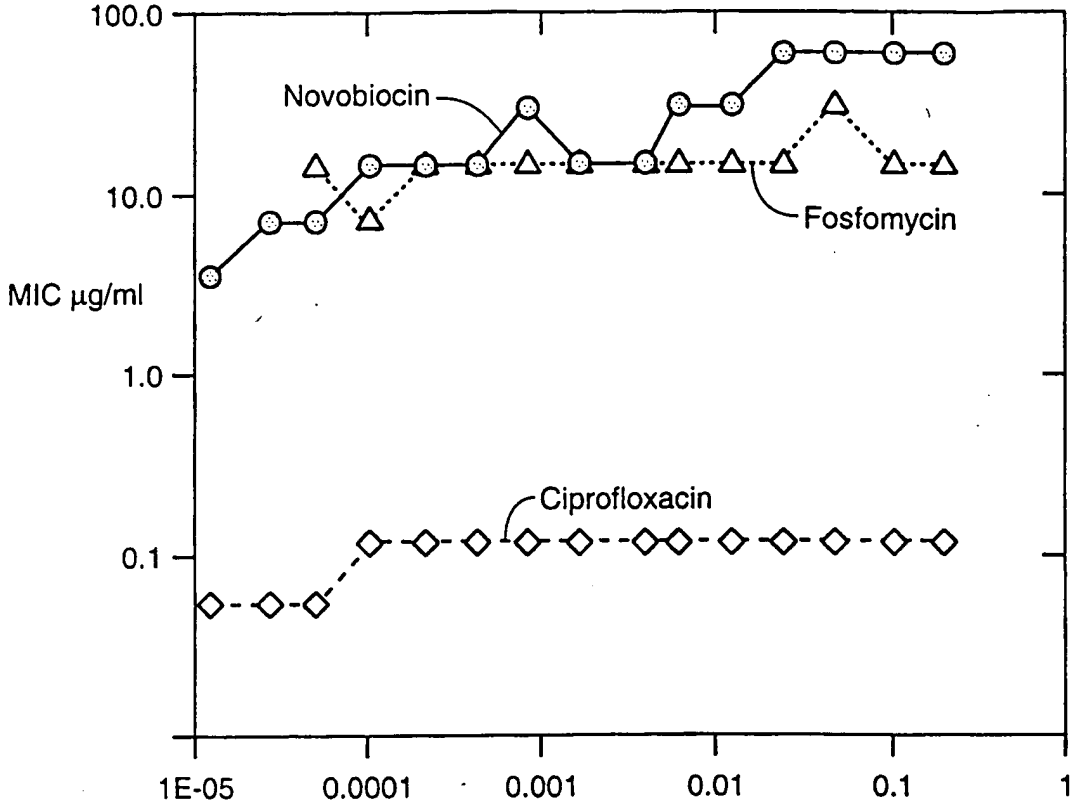
100.0
10.0
1.0
0.1
MIC µg/ml
Novobiocin
Fosfomycin
Ciprofloxacin
1E-05
0.0001
0.001
0.01
0.1
1
% Arabinose

FIG._15

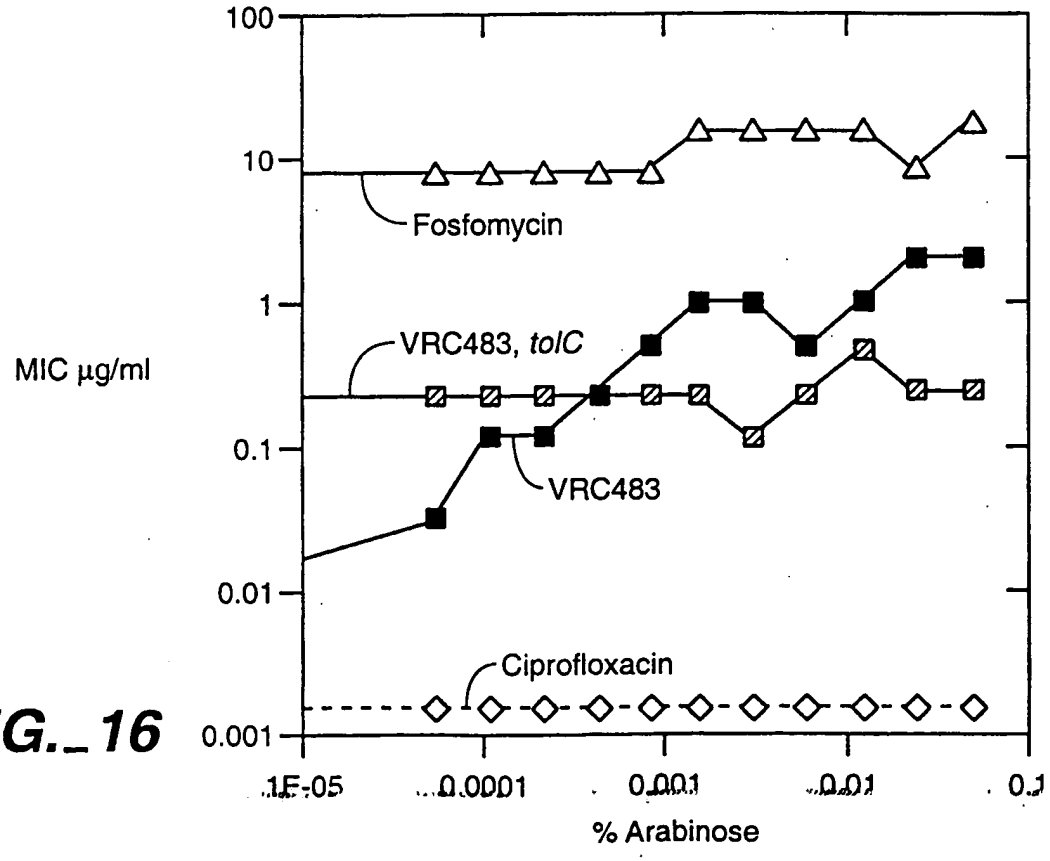
100
10
1
0.1
0.01
0.001
MIC µg/ml
Fosfomycin
VRC483, tolC
VRC483
Ciprofloxacin
1E-05
0.0001
0.001
0.01
0.1
% Arabinose

FIG._16

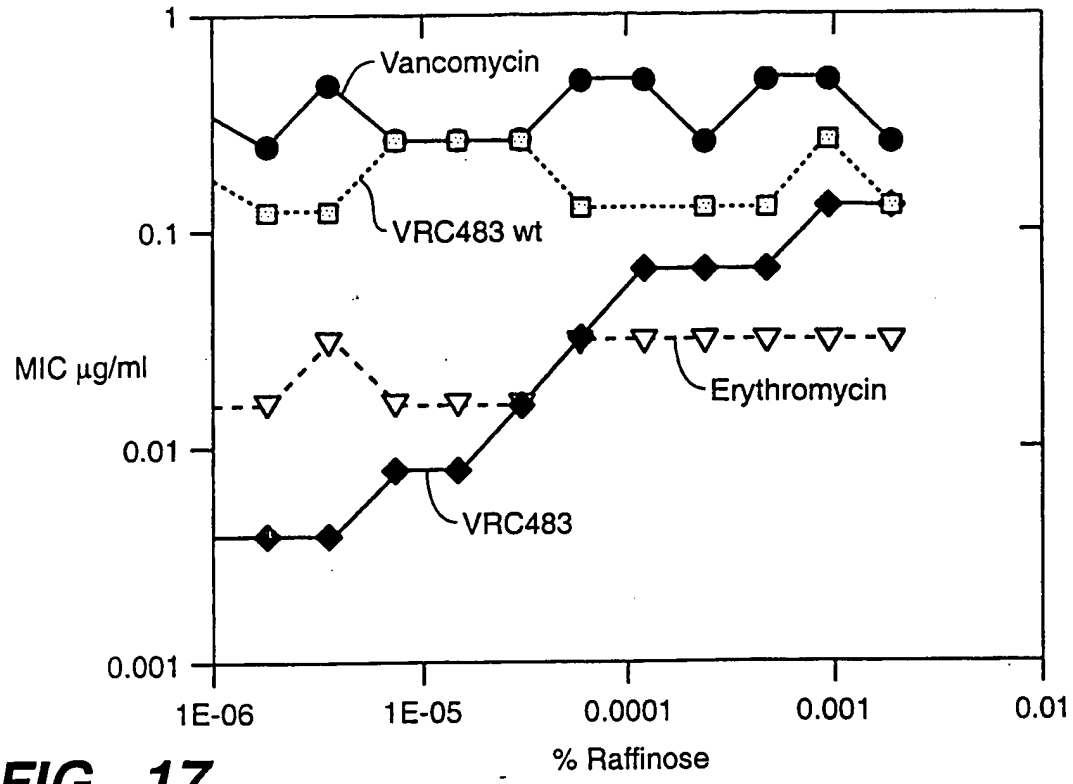
1
0.1
0.01
0.001
MIC µg/ml
Vancomycin
VRC483 wt
Erythromycin
VRC483
1E-06
1E-05
0.0001
0.001
0.01
% Raffinose

FIG._17

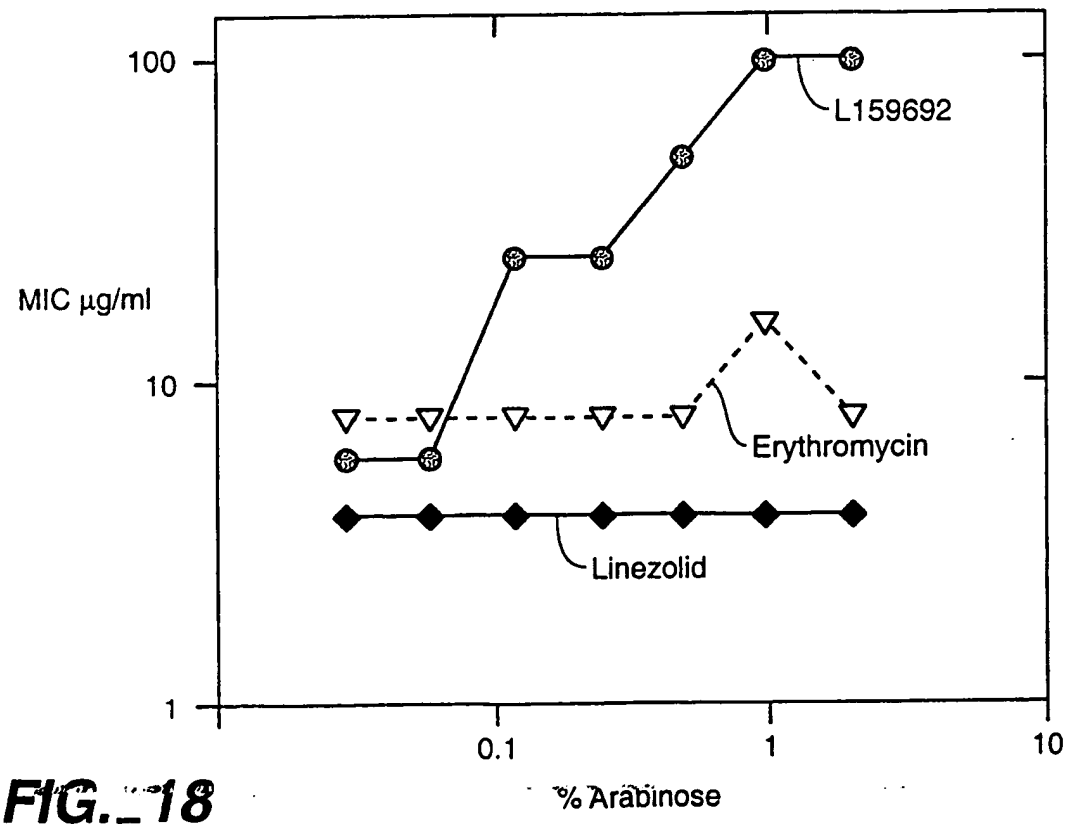
100
10
1
MIC µg/ml
L159692
Erythromycin
Linezolid
0.1
1
10
% Arabinose

FIG._18

SEQUENCE LISTING

<110> Trias, Joaquim
Young, Dennis
Rosenow, Carsten

<120> REGULATED TARGET EXPRESSION FOR SCREENING

<130> 342312000840

<140> Unassigned
<141> Unassigned

<160> 65

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 4065
<212> DNA
<213> S. pneumoniae

<400> 1

attttacttc caactattga gagaaatttc gccactattt agccagattt cttttccgtt 60
atcacattga actaaaagtt ttccattttc tgagatgtct ttagcaagtc ccttgtagtc 120
tttttgctct agtgtgaaag tgacttcttt tcctagaatg aatgactgtt ttttgtatag 180
gtataatagc tcttctgctg gtgtttcgaa gaaagcacgc cagatttcta tgatcaattc 240
attccttgtt ataggagctg tagctttaaa taagctggca gctttttctt ttaattcctg 300
agggaagtct ttaatagtga aattgatcct actccaataa tgatatctgt gactaagcca 360
gtttctacag aggtcattgc ttcagtaagg attcctccaa ttttatgatt gtttagatag 420
atatcattga cccattttat atcgacatct attaaagtta ggttcttaat ggctttgtag 480
acagctccag ctacaagtag tgtgtaggat ggtaatttgt cataggggag atttggttta 540
agatggagtg tcatataaat tccaccttgt ggtgagtaga aggaacgttg aaaacggcct 600
cggcctgctg tttgatagga agctagatag agagtatttg cttcatggcc taaatcaatt 660
gcttcttttg catctagttg tgttgatttt gtttcgggtt taaagctgac tttaattgga 720
agattttctt ctagaatctc tggaagaata aggtcaccat tcattagttt atatcctcta 780
tttttgatac tattaatttc aatgccttct tgttctagtc gcttgatagc tttccaaatt 840
gatgttcggc ttagggatag ttttttctgcg attttttctc cgctgatata gtcggtttct 900
ttagatagga tttggtagac agcttggtag gatttcataa tgttgccttt ctcactagtt 960
ggtattgaga gtattctttt cttgtatgac ttggagactg attaaagtat tgtttataag 1020
ctttcgaaaa atggagtgga tctgaaaaac ctaccgagta tgcaattacc ttgatggact 1080
cttgggtatt ttcgagaagt tgtctagctc ggtgcattcg aacgtagagt aggtattctt 1140
tgggtgataa ggtattaaat tctttgaata cgcttgataa gtagcttctg tgaacggata 1200
gttcttttgc taaatcttga attgtaagtg attgaggata gtggctatca attaatcgtt 1260
tgcattcaag atagagttgg tgggttgatg aaatattctt ttttttctga ttgggagcaa 1320
tagttcccag atgaaacatc agttcatgaa gttgtcccat gatatggagt tgagctaatt 1380
cacttgattt tgtaatctga gcgaagcgga caatgtctga gatgagtttt gcagtagtct 1440
gggtatgaca agtttcagat tggatgagat aggattgatc agaaatttga gaaagagcaa 1500
aataatcagg ggatttccct ccagtgattc ctaaccagta gtaggcccaa ggttctttac 1560
tatctgcttg ataaaaggtt agttcctctg gttttaatag aaagaaatct ccttctttta 1620
aatcaacaat tttacccttg taatgaaatt ttccttgtcc tttactaatg taatgtagga 1680
cgtatgtatc acgaatggct ggaccaaaag agtaattagg tgtgcattcc tcatatccat 1740
aaaagcttag ggcaaggtca attgttccag tctggtattc tgaaaaaact agcatgtgct 1800
acctcctacc taacatttta ccatattctt gagacatttt tctattttgg aaagcgattt 1860
caggtgataa aatataatca ataaagtgat gaggtgaagt aaatgggagt taggatagag 1920
aataatctat tttatgttga gagtaaaaat ctaagtttga ttattgaaaa tcgaaatggc 1980
tacttacttt tgaaacattt aggaaagact attaagaact ataaaggttc caatagtgtt 2040
tatgaacgag accatgcctt ttcaggaaat ccaacggcta ctaatcgaac ctttagttta 2100

gatactcagc gacagatttt tggacaacat ggcttaggag attttaggaa accaactata 2160
caggttcagc atagtgtaac tgaagtaaca gactttcgat ttgtagaagc aaagatttta 2220
aaaggtcaga atggtccaca gggcttacct tctccacata gcatggacga tacagagact 2280
cttgtcttaa tgttagaaga ttctaaggct caacttagtc tgactttgta ttatactact 2340
tttaataatg atgcgactat tgctagctac agtaaattag ataataatag taatcaggaa 2400
gttgtcatcc ataaagattt ttcttttatg gctgattttc cagctgcaga ttacgaaata 2460
gtaactctgc agggtgctta tgctcgtgaa aagactgtta gacgtcaaca ggtagaacaa 2520
ggaatctttt cgattagttc aaaccgaggt gcttctggtc atgctcaaac accagctctt 2580
ctactatgcg aacaaggagt cacagaggat gctgggaatg tgtttgctat tcaactaatg 2640
tatagtggca actttgaagc ttttgttcaa aaaaatcaat tgaatgaagt tcgggtggct 2700
attggcatta atccagaaaa cttttcttgg aagttagccc ctgaggaata ctttgaaaca 2760
ccggtagctt tagtgactca ttcagatcag ggattaactg gtattagtca tgaaagtcag 2820
aattttgtac tgaagcacat tattctaagt gaattttcta aaaaagaacg tccaattcta 2880
atcaataact gggaagctac ttactttgac tttcagagag aaaaactgtt agagttagca 2940
gatgaagcta agaaagttgg cattgaactt tttgtattag atgatggttg gtttggcaat 3000
cgttttgatg ataatcgtgc tttaggtgat tgggttgtta atgaggaaaa actgggtgga 3060
agtctagaaa gtctgatttc agctatccat gaaagaggtt tgcagtttgg actttggtta 3120
gaacccgaaa tgatttctgt agatagtgat ttgtatcgtc aacatcctga ctgggctatt 3180
caggttcctg attatgagca tacttattct cggaatcaat tagtacttaa tcttgccaat 3240
cctcaggtag tagaatactt gaaaagtgtc ttagatcaac tcctatctta tcatgagatt 3300
gattacatta aatgggatat gaaccgcaat atcactaagc tagggaatgg attaacttat 3360
ctagagacac agatgcaatc tcatcagtac atgctgggac tttacgaact cgtttcttat 3420
ctgacagaga agcacagcca tattctcttt gagtcctgct ctggtggtgg tggacgaaat 3480
gatcttggta tgatacgcta tttcccacaa gtctgggcta gtgataatac tgatgccatt 3540
gcacgtttac caattcaata cggttcatcc tatctctatc caaccatttc tatgggggct 3600
catgtgtcag cagtaccgaa tcatcagatg ggacgaatga caccattaga aacacgtggc 3660
catgtagcaa tgatgggaaa tttgggctat gagcttgatt tgacaaattt atcagatgaa 3720
gagaaagcta cgattgctaa tcaggtgaac ttgtataaag aattacgacc agtagttcag 3780
ttaggacagc agtatagact aattaatcct gatactgcat ccaatgaagc agctgtacaa 3840
tttaattacg gaaatcaaac gattgtaacc tacgttcgcg ttttatctgt tgtagagacc 3900
atggaaacaa ctttaaaatt aaaagatttg gatgaagagg gactatataa attacaggaa 3960
aatggcgaag tttactcagg tgcagaacta atgtatgcgg gcttaactgt tatcttatcc 4020
caaggagatt ttttgagtag acagtatatt tttagaaaac tatga 4065

<210> 2
<211> 544
<212> DNA
<213> S. pneumoniae

<400> 2
cataatagca cctcgtgtgt taaaataatg gaacgttgcg tattttgcag acgcaaacgt 60
ttgcgttact tataagtata ctccttttca acgcataatt gcaagcgttt taaaaacttt 120
tgatatttag gagactaact catttagcaa gataactaaa ctatctaatc ctaaggtttg 180
actctcacga gacagtctac aaagtacaaa acctccttag ggagcttgaa tttatatgta 240
acaaagcaca aacgtccata gaaaatagat agggattaga accagggagg tagcccctcc 300
tggtttccct ctttagacag attccatata tttttattaa aaaatactgt attcttatct 360
atttcatact tttcttctat taaaatgcaa aattatgcta tcaaatcaag cttaaaaatt 420
ttttaaaatt tttacaaaaa atacttgcaa ccgttttcta tttgtgctat actaagctat 480
aaaggaaaac gtttgcgttt ccttatcaat aaaacttgct attctttggg aggaatacac 540
tatg 544

<210> 3
<211> 26
<212> DNA
<213> Unknown

<220>
<223> primer

<400> 3

ggccatggat aaatttcgtg ttcagg 26

<210> 4
<211> 27
<212> DNA
<213> Unknown

<400> 4
ggtctagatt attcgccttt cacacgc 27

<210> 5
<211> 28
<212> DNA
<213> Unknown

<400> 5
ggacccgggt ctgatttatc agcgaggc 28

<210> 6
<211> 37
<212> DNA
<213> Unknown

<400> 6
gccatatgtc cggaagctta gtttgttctc agttaac 37

<210> 7
<211> 19
<212> DNA
<213> Unknown

<400> 7
ccggatatgg cgttaaccg 19

<210> 8
<211> 21
<212> DNA
<213> Unknown

<400> 8
cccatggttc cagtaagttc c 21

<210> 9
<211> 18
<212> DNA
<213> Unknown

<400> 9
gtgaatgatg tagccgtc 18

<210> 10
<211> 22
<212> DNA
<213> Unknown

<400> 10
ctcgctaacc aaaccggtaa cc 22

<210> 11
<211> 42

<212> DNA
<213> Unknown

<400> 11
taggttgaat tcatagaaaa tagataggga ttagaaccag gg 42

<210> 12
<211> 72
<212> DNA
<213> Unknown

<400> 12
tgcgaggatc ctacttgtcg tcgtcgtcct tgtagtcgat atcatatgta ttcctcccaa 60
agaatagcaa gt 72

<210> 13
<211> 32
<212> DNA
<213> Unknown

<400> 13
ccatcgcata tgagttcaaa taaactgaca ac 32

<210> 14
<211> 33
<212> DNA
<213> Unknown

<400> 14
cacgacatat gaatcttttt taatcggcaa tcc 33

<210> 15
<211> 33
<212> DNA
<213> Unknown

<400> 15
ctgaatcgat gcagccactt cttctaatat ggc 33

<210> 16
<211> 32
<212> DNA
<213> Unknown

<400> 16
agctatcgat tttctaacct tgtagcctca gc 32

<210> 17
<211> 30
<212> DNA
<213> Unknown

<400> 17
taggttgaat tctcgtgtgt taaaataatg 30

<210> 18
<211> 27
<212> DNA
<213> Unknown

<400> 18
cgcaaacgtt tgcgtttatg agcttag 27

<210> 19
<211> 27
<212> DNA
<213> Unknown

<400> 19
ctaagctcat aaacgcaaac gtttgcg 27

<210> 20
<211> 37
<212> DNA
<213> Unknown

<400> 20
cccggaattc agcttggtag gatttcataa tgttgcc 37

<210> 21
<211> 47
<212> DNA
<213> Unknown

<400> 21
gccgcggatc cgcgcatatg catttacttc acctcatcac tttattg 47

<210> 22
<211> 46
<212> DNA
<213> Unknown

<400> 22
ggggaattcc atatgaattt atttaaaaat ttcttaaaag agtggg 46

<210> 23
<211> 34
<212> DNA
<213> Unknown

<400> 23
gcgctctaga tcattttcgt aacgaatggt gtcg 34

<210> 24
<211> 42
<212> DNA
<213> Unknown

<400> 24
gcgccggaat tccatgtgct acctcctacc taacatttta cc 42

<210> 25
<211> 22
<212> DNA
<213> Unknown

<400> 25
ctcctaccta acattttacc at 22

<210> 26

<211> 27
<212> DNA
<213> Unknown

<400> 26
ttaaaatgtt ccgatacggg tgattgg 27

<210> 27
<211> 32
<212> DNA
<213> Unknown

<400> 27
gcatgccatg gtttcagttt tgcaagtgtt ac 32

<210> 28
<211> 28
<212> DNA
<213> Unknown

<400> 28
cgaagatctt tagttcttat ccttaagc 28

<210> 29
<211> 28
<212> DNA
<213> Unknown

<400> 29
gcggagctcg cagactggca gccagtcg 28

<210> 30
<211> 28
<212> DNA
<213> Unknown

<400> 30
ttggcgcgcc tccagagatg tgttcagg 28

<210> 31
<211> 36
<212> DNA
<213> Unknown

<400> 31
gcatgccatg gcaatcagtc tgattgcggc gttagc 36

<210> 32
<211> 35
<212> DNA
<213> Unknown

<400> 32
cgaagatctt taccgccgct ccagaatctc aaagc 35

<210> 33
<211> 28
<212> DNA
<213> Unknown

<400> 33
gcggagctcg gcgatgccac gcggatgg 28

<210> 34
<211> 32
<212> DNA
<213> Unknown

<400> 34
gcttggcgcg ccaacgagtc cacgctctct cc 32

<210> 35
<211> 51
<212> DNA
<213> Unknown

<400> 35
ggtataccat atgcgagctc caggcgcgcc tgcaggaatt cgatatcaag c 51

<210> 36
<211> 38
<212> DNA
<213> Unknown

<400> 36
tgctctagag ccatatgttc cgctagcttc acgctgcc 38

<210> 37
<211> 41
<212> DNA
<213> Unknown

<400> 37
gcaccggaat tcccgggtca gccagtctaa ctgcgaaagc g 41

<210> 38
<211> 20
<212> DNA
<213> Unknown

<400> 38
cctcgacggt atcgataagc 20

<210> 39
<211> 20
<212> DNA
<213> Unknown

<400> 39
tagcggatcc tacctgacgc 20

<210> 40
<211> 28
<212> DNA
<213> Unknown

<400> 40
cgggatccgc gaagagtacc agtacacc 28

<210> 41

<211> 34
<212> DNA
<213> Unknown

<400> 41
gcatgccatg gcatcgaatt cttatgactc ctcc 34

<210> 42
<211> 30
<212> DNA
<213> Unknown

<400> 42
gtccgatcgt taaatatcga tattcgccgc 30

<210> 43
<211> 28
<212> DNA
<213> Unknown

<400> 43
gcggagctca gcgattgctc aagcagcg 28

<210> 44
<211> 33
<212> DNA
<213> Unknown

<400> 44
gcttggcgcg cctctcgctc atttatactt ggg 33

<210> 45
<211> 32
<212> DNA
<213> Unknown

<400> 45
tcagcggccg ccagcgtgca gattgaagat gc 32

<210> 46
<211> 32
<212> DNA
<213> Unknown

<400> 46
tgactcgagc cgtgtagtag ctgatatcac gg 32

<210> 47
<211> 20
<212> DNA
<213> Unknown

<400> 47
ccaccataat tgacgaacgc 20

<210> 48
<211> 21
<212> DNA
<213> Unknown

<400> 48
gtcttcggta cggtcatggt g 21

<210> 49
<211> 35
<212> DNA
<213> Unknown

<400> 49
ggggaattcc atatgtctgc aatagaacgt attac 35

<210> 50
<211> 30
<212> DNA
<213> Unknown

<400> 50
ccgcggatcc aaatcgtagg cttcctgtgg 30

<210> 51
<211> 40
<212> DNA
<213> Unknown

<400> 51
ggcgcggatc cttaagcttc gatttctgta accatacctg 40

<210> 52
<211> 28
<212> DNA
<213> Unknown

<400> 52
tcggatccgg ctacgcaatg atgggttc 28

<210> 53
<211> 41
<212> DNA
<213> Unknown

<400> 53
cccatccact aaacttaaac atgtcctttg tttgatcatc g 41

<210> 54
<211> 41
<212> DNA
<213> Unknown

<400> 54
tgtttaagtt tagtggatgg gttggccttc aaagcgcctt c 41

<210> 55
<211> 27
<212> DNA
<213> Unknown

<400> 55
gtggatccgt aatgcaagat cttgcgc 27

<210> 56

<211> 35
<212> DNA
<213> Unknown

<400> 56
ggttccatgg caatcaaaca aaggacactt aaacg 35

<210> 57
<211> 30
<212> DNA
<213> Unknown

<400> 57
gtcagatctt tatgccagta cagctgaagg 30

<210> 58
<211> 26
<212> DNA
<213> Unknown

<400> 58
gacccgggtg ataccattcg cgagcc 26

<210> 59
<211> 28
<212> DNA
<213> Unknown

<400> 59
gagtcgacgc agcgtttgct gcatatcc 28

<210> 60
<211> 21
<212> DNA
<213> Unknown

<400> 60
gctcaactta gtctgacttt g 21

<210> 61
<211> 24
<212> DNA
<213> Unknown

<400> 61
caaacacatt cccagcatcc tctg 24

<210> 62
<211> 30
<212> DNA
<213> Unknown

<400> 62
cgcggatcct cgagaagttg tctagctcgg 30

<210> 63
<211> 29
<212> DNA
<213> Unknown

<400> 63
ccggaattct aggaatcact ggagggaaa 29

<210> 64
<211> 33
<212> DNA
<213> Unknown

<400> 64
ccgcggatcc gctacaagta gtgtgtagga tgg 33

<210> 65
<211> 31
<212> DNA
<213> Unknown

<400> 65
gccggaattc aatcctacca agctgtctac c 31

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/08164

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :435/6, 29, 34, 69.1, 252.1, 253.1, 254.1, 255.1, 320.1, 325

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 29, 34, 69.1, 252.1, 253.1, 254.1, 255.1, 320.1, 325

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GHRAYEB et al. Secretion Cloning Vectors in E. coli. The EMBO Journal. 1984, Vol. 3, No. 10, pages 2437-2442, see entire document.	1-77
A	DEL CASTILLO et al. An Unusual Mechanism for Resistance to the Antibiotic Coumermycin A1. Proc. Natl. Acad. Sci. October 1991, Vol. 88, pages 8860-8864, see the entire document.	1-77
A	CHENG et al. Isolation of Gram Quantities of EcoRI Restriction and Modification Enzymes from an Overproducing Strain. The Journal of Biological Chemistry. 25 September 1984, Vol. 259, No. 18, pages 11571-11575, see the entire document.	1-77

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 03 JULY 1999	Date of mailing of the international search report 24 AUG 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer NANCY J. DEGEN Telephone No. (703) 308-0196

Form PCT ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/08164

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C07H 21/04; C12N 1/12, 1/14, 1/20, 15/63; C12P 21/02; C12Q 1/00, 1/02, 1/04

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, BIOSIS, EMBASE, MEDLINE, DERWENT
search terms: cell? express? basal? polypeptide? protein? ma? dna? untranslat? enzym? transcri? pbad? control? essential? arabinose? ribose?